GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2006, 00:28:29 ; Search time 46 Seconds
(without alignments)
1365.946 Million cell updates/sec

Title: US-10-681-659-1
Perfect score: 4086
Sequence: 1 MKTWVKIVFGVATSAVLALL.....TNHLYTHMTHFLKQCFSLSD 760

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4069	99.6	760	1	US-08-230-491A-2	Sequence 2, Appli
2	4069	99.6	760	1	US-08-619-280A-2	Sequence 2, Appli
3	4069	99.6	760	1	US-08-940-391-2	Sequence 2, Appli
4	4069	99.6	760	2	US-09-265-606-2	Sequence 2, Appli
5	3705.5	90.7	761	2	US-09-518-550-26	Sequence 26, Appl
6	2823	69.1	547	2	US-09-949-016-8330	Sequence 8330, Ap
7	2190.5	53.6	766	1	US-08-230-491A-3	Sequence 3, Appli
8	2190.5	53.6	766	1	US-08-619-280A-3	Sequence 3, Appli
9	2190.5	53.6	766	1	US-08-940-391-3	Sequence 3, Appli
10	2190.5	53.6	766	2	US-09-794-236-1	Sequence 1, Appli
11	2190.5	53.6	766	2	US-09-265-606-3	Sequence 3, Appli
12	2190.5	53.6	775	2	US-09-949-016-10450	Sequence 10450, A
13	2187.5	53.5	766	2	US-10-002-593-6	Sequence 6, Appli
14	2187.5	53.5	766	2	US-09-949-016-6146	Sequence 6146, Ap
15	2187.5	53.5	766	2	US-10-423-714-6	Sequence 6, Appli
16	2185.5	53.5	766	2	US-09-518-550-27	Sequence 27, Appl
17	2180	53.4	755	4	PCT-US93-07923-3	Sequence 3, Appli
18	2178	53.3	759	4	PCT-US93-07923-2	Sequence 2, Appli
19	1489	36.4	593	4	PCT-US93-07923-11	Sequence 11, Appl
20	1223.5	29.9	796	2	US-09-976-674-5	Sequence 5, Appli
21	1168	28.6	706	2	US-09-976-674-41	Sequence 41, Appl
22	1168	28.6	743	2	US-10-363-937-4	Sequence 4, Appli
23	1146.5	28.1	691	2	US-09-976-674-43	Sequence 43, Appl
24	986.5	24.1	771	2	US-09-079-592-2	Sequence 2, Appli
25	949.5	23.2	771	2	US-09-462-284-2	Sequence 2, Appli
26	895	21.9	723	2	US-09-518-550-29	Sequence 29, Appl
27	880	21.5	676	2	US-09-518-550-42	Sequence 42, Appl
28	841	20.6	710	2	US-09-518-550-28	Sequence 28, Appl
29	814.5	19.9	818	2	US-09-462-845-3	Sequence 3, Appli
30	814.5	19.9	818	2	US-10-402-312-3	Sequence 3, Appli
31	814.5	19.9	818	2	US-10-401-437-3	Sequence 3, Appli
32	814.5	19.9	818	2	US-10-402-067-3	Sequence 3, Appli
33	814.5	19.9	818	2	US-10-401-436-3	Sequence 3, Appli
34	721	17.6	931	2	US-09-079-592-11	Sequence 11, Appl
35	610.5	14.9	323	2	US-09-270-767-45296	Sequence 45296, A
36	576	14.1	494	2	US-09-248-796A-19472	Sequence 19472, A
37	522.5	12.8	732	2	US-09-518-550-30	Sequence 30, Appl
38	492.5	12.1	981	2	US-09-902-540-16812	Sequence 16812, A
39	476.5	11.7	526	2	US-09-248-796A-19924	Sequence 19924, A
40	456	11.2	882	2	US-09-976-674-1	Sequence 1, Appli
41	456	11.2	882	2	US-10-070-464-1	Sequence 1, Appli
42	425.5	10.4	879	2	US-09-976-674-33	Sequence 33, Appl
43	425.5	10.4	879	2	US-09-976-674-35	Sequence 35, Appl
44	425	10.4	863	2	US-09-976-674-3	Sequence 3, Appli
45	425	10.4	892	2	US-09-976-674-23	Sequence 23, Appl

ALIGNMENTS

```
RESULT 1
US-08-230-491A-2
; Sequence 2, Application US/08230491A
; Patent No. 5587299
;  GENERAL INFORMATION:
;    APPLICANT:  Rettig, Wolfgang J.; Scanlan, Matthew J.;
;    APPLICANT:  Garin-Chesa, Pilar; Old, Lloyd J.
;    TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULE CODING FOR
;    TITLE OF INVENTION:  FIBROBLAST ACTIVATION PROTEIN ' AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FELFE & LYNCH
;      STREET:  805 THIRD AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE 3.5 inch 1.2 MB STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT - ASC II
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/230,491A
;      FILING DATE:  20-APRIL-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5587299man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 330
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  760 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-230-491A-2

  Query Match             99.6%;  Score 4069;  DB 1;  Length 760;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches  757;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
```

```
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              |||||||||||||||||||||||| |||||||||||||||||||| ||||||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760


RESULT 2
US-08-619-280A-2
; Sequence 2, Application US/08619280A
; Patent No. 5767242
;  GENERAL INFORMATION:
;    APPLICANT:  Zimmermann, Rainer; Park, John E.;
;    APPLICANT:  Rettig, Wolfgang; Old, Lloyd J.
;    TITLE OF INVENTION:  ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
;    TITLE OF INVENTION:  ALPHA, AND USES THEREOF
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 2.0 MB  storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,280A
;      FILING DATE:  18-MARCH-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/230,491
;      FILING DATE:  20-APRIL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5767242man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5330.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  760 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-619-280A-2

  Query Match             99.6%;  Score 4069;  DB 1;  Length 760;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches  757;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              |||||||||||||||||||||||| |||||||||||||||||||| ||||||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
```

```
RESULT 3
US-08-940-391-2
; Sequence 2, Application US/08940391
; Patent No. 5965373
;  GENERAL INFORMATION:
;    APPLICANT:  Zimmermann, Rainer; Park, John E.;
;    APPLICANT:  Rettig, Wolfgang; Old, Lloyd J.
;    TITLE OF INVENTION:  ISOLATED DIMERIC FIBROBLAST ACTIVATION
;    TITLE OF INVENTION:  PROTEIN ALPHA, AND USES THEREOF
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 2.0 MB  storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,391
;      FILING DATE:  01-OCT-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/619,280
;      FILING DATE:  18-MARCH-1996
;      APPLICATION NUMBER:  08/230,491
;      FILING DATE:  20-APRIL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5965373man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5330.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  760 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-940-391-2

  Query Match             99.6%;  Score 4069;  DB 1;  Length 760;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches  757;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              |||||||||||||||||||||||| |||||||||||||||||| ||||||||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||

Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760


RESULT 4
US-09-265-606-2
; Sequence 2, Application US/09265606
; Patent No. 6846910
;  GENERAL INFORMATION:
;    APPLICANT:  Zimmermann, Rainer; Park, John E.;
;    APPLICANT:  Rettig, Wolfgang; Old, Lloyd J.
;    TITLE OF INVENTION:  ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
;    TITLE OF INVENTION:  ALPHA, AND USES THEREOF
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 2.0 MB  storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/265,606
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,280
;      FILING DATE:  18-MARCH-1996
;      APPLICATION NUMBER:  08/230,491
;      FILING DATE:  20-APRIL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 6846910man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5330.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  760 amino acids
;      TYPE:  amino acid
```

```
;      TOPOLOGY:  linear
US-09-265-606-2

  Query Match             99.6%;  Score 4069;  DB 2;  Length 760;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches  757;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              |||||||||||||||||||||||| |||||||||||||||||| ||||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760


RESULT 5
US-09-518-550-26
; Sequence 26, Application US/09518550
; Patent No. 6875851
; GENERAL INFORMATION:
;  APPLICANT: TRAVIS, James
;  APPLICANT:  POTEMPA, Jan
;  APPLICANT:  BANBULA, Agnieszka
;  TITLE OF INVENTION: PROLYL PEPTIDASES AND METHODS OF USE
;  FILE REFERENCE: 235.00190101
;  CURRENT APPLICATION NUMBER: US/09/518,550
;  CURRENT FILING DATE:  2000-03-03
;  PRIOR APPLICATION NUMBER: 60/123,148
;  PRIOR FILING DATE: 1999-03-05
;  PRIOR APPLICATION NUMBER: PCT/US00/05551
;  PRIOR FILING DATE: 2000-03-03
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
;   LENGTH: 761
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-518-550-26

  Query Match             90.7%;  Score 3705.5;  DB 2;  Length 761;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches  681;  Conservative   37;  Mismatches   42;  Indels    1;  Gaps    1;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||:| |||| | | |||:|:||||||||:  | || ||||||||||||||||:|||
Db          1 MKTWLKTVFGVTTLAALALVVICIVLRPSRVYKPEGNTKRALTLKDILNGTFSYKTYFPN 60

Qy         61 WISGQEYLHQSADNNIVFYNIETRESYIILSNSTMKSVNATDYGLSPDRQFVYLESDYSK 120
              ||| ||||||| |:||| ||||| :|| |||| |||||||::||||||||||||||||||
Db         61 WISEQEYLHQSEDDNIVFYNIETRESYIILSNSTMKSVNATDYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              |||||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLQNGEFVRGYELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              ||||: ||||:|||||||||||||||||||||||:|||||| ||||:|||:|||||||
Db        181 FQITYTGRENRIFNGIPDWVYEEEMLATKYALWWSPDGKFLAYVEFNDSDIPIIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              | |||||||||||||||||||:||:||||| :||| ||||| ||||||||||||||||:
Db        241 DGQYPRTINIPYPKAGAKNPVVRVFIVDTTYPHHVGPMEVPVPEMIASSDYYFSWLTWVS 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||  |:||| |||:|||||||||||||||| || |
Db        301 SERVCLQWLKRVQNVSVLSICDFREDWHAWECPKNQEHVEESRTGWAGGFFVSTPAFSQD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              | |||||||||||||||||||||||||||||||||| |||||||||||||||| |||
Db        361 ATSYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAIYIFRVTQDSLFYSSNEFEGYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||: |||||||||||||||||||||||  |||||||||||:|||||||||
Db        421 RRNIYRISIGNSPPSKKCVTCHLRKERCQYYTASFSYKAKYYALVCYGPGLPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              |||||::|||||||||:|:|||||| |||||:   :| ||||||||||||||||||||
Db        481 TDQEIQVLEENKELENSLRNIQLPKVEIKKLKDGGLTFWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              |||||||||:|||||||:|||||||||||||||||||||||| |:|||||||||||||||
Db        541 YGGPCSQSVKSVFAVNWITYLASKEGIVIALVDGRGTAFQGDKFLHAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              :|||||||||||||:|||||||||||||||||||||||||||||||||||||||||||:|
Db        601 LTAVRKFIEMGFIDEERIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASIY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 SERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||: || | ||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGISSGRSQNHLYTHMTHFLKQCFSLSD 761


RESULT 6
US-09-949-016-8330
; Sequence 8330, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8330
;   LENGTH: 547
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-8330

  Query Match             69.1%;  Score 2823;  DB 2;  Length 547;
  Best Local Similarity   97.0%;  Pred. No. 5.9e-248;
  Matches  525;  Conservative    0;  Mismatches    0;  Indels   16;  Gaps    1;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 67

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 127

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 187

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 247

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 307

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 367

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 427

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCY---------- 474

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 ---EIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 531

Qy        541 Y 541
              |
Db        532 Y 532


RESULT 7
US-08-230-491A-3
; Sequence 3, Application US/08230491A
; Patent No. 5587299
;  GENERAL INFORMATION:
;    APPLICANT:  Rettig, Wolfgang J.; Scanlan, Matthew J.;
;    APPLICANT:  Garin-Chesa, Pilar; Old, Lloyd J.
;    TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULE CODING FOR
;    TITLE OF INVENTION:  FIBROBLAST ACTIVATION PROTEIN ' AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FELFE & LYNCH
;      STREET:             805 THIRD AVENUE
;      CITY:  NEW YORK
;      STATE:              NEW YORK
;      COUNTRY:            USA
;      ZIP:                10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE 3.5 inch 1.2 MB STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT - ASC II
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/230,491A
;      FILING DATE:  20-APRIL-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5587299man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 330
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  766 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-230-491A-3

  Query Match             53.6%;  Score 2190.5;  DB 1;  Length 766;
  Best Local Similarity   51.5%;  Pred. No. 4e-190;
  Matches  395;  Conservative  149;  Mismatches  212;  Indels   11;  Gaps    7;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  |:: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db          1 MKTPWKVLLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
              |||  |||::  :|||:::| | | |   | | |       |  :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
              | ::||: |:|: |:||| |||||||: :   |||||||| |||||:||||::|:| ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :|||||| || |: |:::|   : |      ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
              :|| | ||: ||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||::  || ||:|:| :  |      |::| |  |||||: |||  |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||    |: :::||:|  |:  |:|:|:| :::  : ::|   ||:|||| ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538
```

```
Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||| | |:| :|||:| |||||||||||||:|: | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| ||||||:||||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |:|||:  |||||||:|::||: |  :  |:||||:||:||||||
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765


RESULT 8
US-08-619-280A-3
; Sequence 3, Application US/08619280A
; Patent No. 5767242
;  GENERAL INFORMATION:
;    APPLICANT:  Zimmermann, Rainer; Park, John E.;
;    APPLICANT:  Rettig, Wolfgang; Old, Lloyd J.
;    TITLE OF INVENTION:  ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
;    TITLE OF INVENTION:  ALPHA, AND USES THEREOF
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 2.0 MB  storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,280A
;      FILING DATE:  18-MARCH-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/230,491
;      FILING DATE:  20-APRIL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5767242man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5330.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  766 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-619-280A-3

  Query Match             53.6%;  Score 2190.5;  DB 1;  Length 766;
  Best Local Similarity   51.5%;  Pred. No. 4e-190;
  Matches  395;  Conservative  149;  Mismatches  212;  Indels   11;  Gaps    7;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  |:: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db          1 MKTPWKVLLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
                |||  |||::  :|||:::| | | |   | | |      |  :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
              | ::||: |:|: |:||| |||||||: :   |||||||| |||||:||||::|:| ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :||||||| || |: |:::|   : |      ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
                :|| | ||: ||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||::  || ||:|:| :  |     |::| |  |||||: ||| |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||    |: :::||:|  |:  |:|:|:| :::  : ::|   ||:|||| ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||| | |:| :|||:| |||||||||||||:|: | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| ||||||:||||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |:|||:  |||||||:|::||: |  :  |:||||:||:||||||
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765


RESULT 9
US-08-940-391-3
; Sequence 3, Application US/08940391
; Patent No. 5965373
;  GENERAL INFORMATION:
;    APPLICANT:  Zimmermann, Rainer; Park, John E.;
;    APPLICANT:  Rettig, Wolfgang; Old, Lloyd J.
;    TITLE OF INVENTION:  ISOLATED DIMERIC FIBROBLAST ACTIVATION
;    TITLE OF INVENTION:  PROTEIN ALPHA, AND USES THEREOF
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 2.0 MB  storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,391
;      FILING DATE:  01-OCT-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/619,280
;      FILING DATE:  18-MARCH-1996
```

```
;      APPLICATION NUMBER:  08/230,491
;      FILING DATE:  20-APRIL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5965373man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5330.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  766 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-940-391-3

  Query Match             53.6%;  Score 2190.5;  DB 1;  Length 766;
  Best Local Similarity   51.5%;  Pred. No. 4e-190;
  Matches  395;  Conservative  149;  Mismatches  212;  Indels   11;  Gaps    7;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  |:: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db          1 MKTPWKVLLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
               ||.|  |||::  :|||:::| | | |   | | |      |  :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | |.||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
               | ::||: |:|: |:||| |||||||: :   |||||||| |||||:||||::|:| ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :|||||||  || |: |:::|   : |      ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
                :|| | ||: ||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||::  || ||:|:| :  |      |::| |  |||||: |||  |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||    |: :::||:|  |:  |:|:|:| :::  : ::|   ||:|||| ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||  | |:| :|||:| ||||||||||||||:| | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| ||||||:|||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |:||||:  ||||||||:|::||: |  :  |:||||:||:||||||
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765
```

```
RESULT 10
US-09-794-236-1
; Sequence 1, Application US/09794236
; Patent No. 6337069
; GENERAL INFORMATION:
;  APPLICANT: Grouzmann, Eric
;  APPLICANT:  Lacroix, Jean-Silvain
;  APPLICANT:  Monod, Michel
;  TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
;  FILE REFERENCE: 81985/276823
;  CURRENT APPLICATION NUMBER: US/09/794,236
;  CURRENT FILING DATE:  2001-02-28
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 766
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-794-236-1

  Query Match             53.6%;  Score 2190.5;  DB 2;  Length 766;
  Best Local Similarity   51.5%;  Pred. No. 4e-190;
  Matches  395;  Conservative  149;  Mismatches  212;  Indels   11;  Gaps    7;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  |:: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db          1 MKTPWKVLLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
               |||  |||::  :|||:::| | | |   | | |      |  :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
               | ::||: |:|: |:||| |||||||: :   |||||||| |||||:||||::|:| ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :|||||||  || |: |:::|   : |      ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
                :|| | ||: ||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||::  || ||:|:| :  |      |::| |  |||||: |||  |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||    |: :::||:|  |:  |:|:|:| :::  : ::|   ||:|||| ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||  | |:| :|||:| ||||||||||||||:| | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| ||||||:|||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
```

```
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765


RESULT 11
US-09-265-606-3
; Sequence 3, Application US/09265606
; Patent No. 6846910
;  GENERAL INFORMATION:
;    APPLICANT:  Zimmermann, Rainer; Park, John E.;
;    APPLICANT:  Rettig, Wolfgang; Old, Lloyd J.
;    TITLE OF INVENTION:  ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
;    TITLE OF INVENTION:  ALPHA, AND USES THEREOF
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 2.0 MB  storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/265,606
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,280
;      FILING DATE:  18-MARCH-1996
;      APPLICATION NUMBER:  08/230,491
;      FILING DATE:  20-APRIL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 6846910man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5330.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  766 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-09-265-606-3

  Query Match             53.6%;  Score 2190.5;  DB 2;  Length 766;
  Best Local Similarity   51.5%;  Pred. No. 4e-190;
  Matches  395;  Conservative  149;  Mismatches  212;  Indels   11;  Gaps    7;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  |:: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db          1 MKTPWKVLLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
               |||  |||::  :|||:::| | | |   | | |      |  :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
               | ::||: |:|: |:||| ||||||: :   |||||||| |||||:||||::|:| ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :||||||| || |: |:::|   : |      ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
               :|| | ||: |||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||::  || ||:|:| :  |      |::| |  ||||||: |||  |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||    |: :::||:|  |:  |:|:|:| :::  : ::|   ||:|||| ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||| | |:| :|||:| ||||||||||||||:|: | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| ||||||:||||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |:||||:  |||||||:|::||: |   :  |:||||:||:|||||
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765


RESULT 12
US-09-949-016-10450
; Sequence 10450, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10450
;   LENGTH: 775
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-10450

  Query Match             53.6%;  Score 2190.5;  DB 2;  Length 775;
  Best Local Similarity   51.5%;  Pred. No. 4.1e-190;
  Matches  395;  Conservative  149;  Mismatches  212;  Indels   11;  Gaps    7;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  |:: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db         10 MKTPWKVLLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 69

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
               |||  |||::  :|||:::| | | |   | | |      |  :| :||| ||: || :
Db         70 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 128

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        129 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 188
```

```
Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
Db        189 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 248

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
Db        249 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 308

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
Db        309 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 368

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
Db        369 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 428

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
Db        429 NEYKGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 487

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
Db        488 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 547

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
Db        548 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 607

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
Db        608 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 667

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
Db        668 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 727

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
Db        728 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 774


RESULT 13
US-10-002-593-6
; Sequence 6, Application US/10002593
; Patent No. 6586198
; GENERAL INFORMATION:
;  APPLICANT: Vanderbilt University
;  APPLICANT:  Brown, Nancy J.
;  TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTIN
;  TITLE OF INVENTION:  INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
;  FILE REFERENCE: Atty Docket No. 6586198 1242/48/2
;  CURRENT APPLICATION NUMBER: US/10/002,593
;  CURRENT FILING DATE:  2001-10-31
;  PRIOR APPLICATION NUMBER: 60/244,524
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 766
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-002-593-6

  Query Match             53.5%;  Score 2187.5;  DB 2;  Length 766;
  Best Local Similarity   51.6%;  Pred. No. 7.6e-190;
  Matches  396;  Conservative  148;  Mismatches  212;  Indels   11;  Gaps    7;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
Db          1 MKTPWKILLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
Db        420 NEYKGMPGGRNLYKIQLIDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765


RESULT 14
US-09-949-016-6146
; Sequence 6146, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6146
;   LENGTH: 766
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-6146
```

```
  Query Match             53.5%;  Score 2187.5;  DB 2;  Length 766;
  Best Local Similarity   51.6%;  Pred. No. 7.6e-190;
  Matches  396;  Conservative  148;  Mismatches  212;  Indels   11;  Gaps    7;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  ||: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db          1 MKTPWKILLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
               |||  |||::  :|||:::| | | |   | | |       |  :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
               | ::||: |:|: |:||| |||||||: :   |||||||| |||||:||||::|:| ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :|||||| || |: |:::|   : |      ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
              :|| | ||: ||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||::  || ||:|:| :  |     |::| |  ||||||: |||  |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLIDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||    |: :::||:|  |:  |:|:|:| :::  : ::|   ||:|||| ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||| | |:| :|||:| |||||||||||||:|: | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| |||||:|||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |:|||:  |||||||:|::||: |  :  |:||||:||:||||||
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765


RESULT 15
US-10-423-714-6
; Sequence 6, Application US/10423714
; Patent No. 6887679
; GENERAL INFORMATION:
;  APPLICANT: Vanderbilt University
;  APPLICANT:  Brown, Nancy J.
;  TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTIN
;  TITLE OF INVENTION:  INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
;  FILE REFERENCE: Atty Docket No. 6887679 1242/48/2
;  CURRENT APPLICATION NUMBER: US/10/423,714
;  CURRENT FILING DATE:  2003-04-25
;  PRIOR APPLICATION NUMBER: 60/244,524
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 6
;   LENGTH: 766
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-423-714-6

  Query Match             53.5%;  Score 2187.5;  DB 2;  Length 766;
  Best Local Similarity   51.6%;  Pred. No. 7.6e-190;
  Matches  396;  Conservative  148;  Mismatches  212;  Indels   11;  Gaps    7;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  ||: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db          1 MKTPWKILLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
               |||  |||::  :|||:::| | | |   | | |       |  :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
               | ::||: |:|: |:||| |||||||: :   |||||||| |||||:||||::|:| ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :|||||| || |: |:::|   : |      ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
              :|| | ||: ||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||::  || ||:|:| :  |     |::| |  ||||||: |||  |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLIDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||    |: :::||:|  |:  |:|:|:| :::  : ::|   ||:|||| ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||| | |:| :|||:| |||||||||||||:|: | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| |||||:|||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |:|||:  |||||||:|::||: |  :  |:||||:||:||||||
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765


Search completed: January 28, 2006, 00:36:14
Job time : 48 secs
```

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

Withdrawn from Queen's University Library

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2006, 00:34:31 ; Search time 116 Seconds
(without alignments)
2737.502 Million cell updates/sec

Title: US-10-681-659-1
Perfect score: 4086
Sequence: 1 MKTWVKIVFGVATSAVLALL.....TNHLYTHMTHFLKQCFSLSD 760

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4086	100.0	760	5	US-10-723-860-4171	Sequence 4171, Ap
2	4069	99.6	760	3	US-09-265-606-2	Sequence 2, Appli
3	4069	99.6	760	4	US-10-177-293-136	Sequence 136, App
4	4069	99.6	760	4	US-10-301-822-55	Sequence 55, Appl
5	4069	99.6	760	5	US-10-884-070A-13	Sequence 13, Appl
6	2204	53.9	767	4	US-10-165-603-4	Sequence 4, Appli
7	2204	53.9	767	4	US-10-794-899-38	Sequence 38, Appl
8	2199	53.8	767	5	US-10-770-712-2	Sequence 2, Appli
9	2191.5	53.6	766	4	US-10-295-027-590	Sequence 590, App
10	2190.5	53.6	766	3	US-09-265-606-3	Sequence 3, Appli
11	2190.5	53.6	766	3	US-09-993-959-1	Sequence 1, Appli
12	2190.5	53.6	766	5	US-10-952-459-18	Sequence 18, Appl
13	2190.5	53.6	766	5	US-10-631-467-565	Sequence 565, App
14	2187.5	53.5	766	4	US-10-002-593-6	Sequence 6, Appli
15	2187.5	53.5	766	4	US-10-165-603-7	Sequence 7, Appli
16	2187.5	53.5	766	4	US-10-423-714-6	Sequence 6, Appli
17	2187.5	53.5	766	4	US-10-295-027-922	Sequence 922, App
18	2187.5	53.5	766	4	US-10-794-899-41	Sequence 41, Appl
19	2187.5	53.5	766	6	US-11-041-674-6	Sequence 6, Appli
20	2185.5	53.5	766	5	US-10-476-264-147	Sequence 147, App
21	2185.5	53.5	766	5	US-10-770-712-1	Sequence 1, Appli
22	2183.5	53.4	760	5	US-10-631-467-1390	Sequence 1390, Ap
23	2175.5	53.2	760	5	US-10-770-712-3	Sequence 3, Appli
24	1301	31.8	504	4	US-10-072-012-863	Sequence 863, App
25	1223.5	29.9	789	5	US-10-476-264-71	Sequence 71, Appl
26	1223.5	29.9	796	3	US-09-976-674-5	Sequence 5, Appli
27	1223.5	29.9	796	3	US-09-870-133-2	Sequence 2, Appli
28	1223.5	29.9	796	4	US-10-160-501-5	Sequence 5, Appli
29	1223.5	29.9	796	5	US-10-982-512-5	Sequence 5, Appli
30	1223.5	29.9	796	5	US-10-433-757-2	Sequence 2, Appli
31	1223.5	29.9	796	5	US-10-476-264-68	Sequence 68, Appl
32	1223.5	29.9	796	5	US-10-476-264-106	Sequence 106, App
33	1223.5	29.9	796	5	US-10-476-264-110	Sequence 110, App
34	1223.5	29.9	796	5	US-10-476-264-142	Sequence 142, App
35	1223.5	29.9	796	5	US-10-476-264-146	Sequence 146, App
36	1223.5	29.9	797	5	US-10-476-264-36	Sequence 36, Appl
37	1221.5	29.9	798	4	US-10-210-130-120	Sequence 120, App
38	1206.5	29.5	746	5	US-10-476-264-69	Sequence 69, Appl
39	1206.5	29.5	746	5	US-10-476-264-70	Sequence 70, Appl
40	1206.5	29.5	746	5	US-10-476-264-72	Sequence 72, Appl
41	1204	29.5	228	5	US-10-476-264-115	Sequence 115, App
42	1168	28.6	706	3	US-09-976-674-41	Sequence 41, Appl
43	1168	28.6	706	5	US-10-982-512-41	Sequence 41, Appl
44	1168	28.6	743	4	US-10-363-937-4	Sequence 4, Appli
45	1168	28.6	743	6	US-11-167-040-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-10-723-860-4171
; Sequence 4171, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4171
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-4171

Query Match 100.0%; Score 4086; DB 5; Length 760;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
```

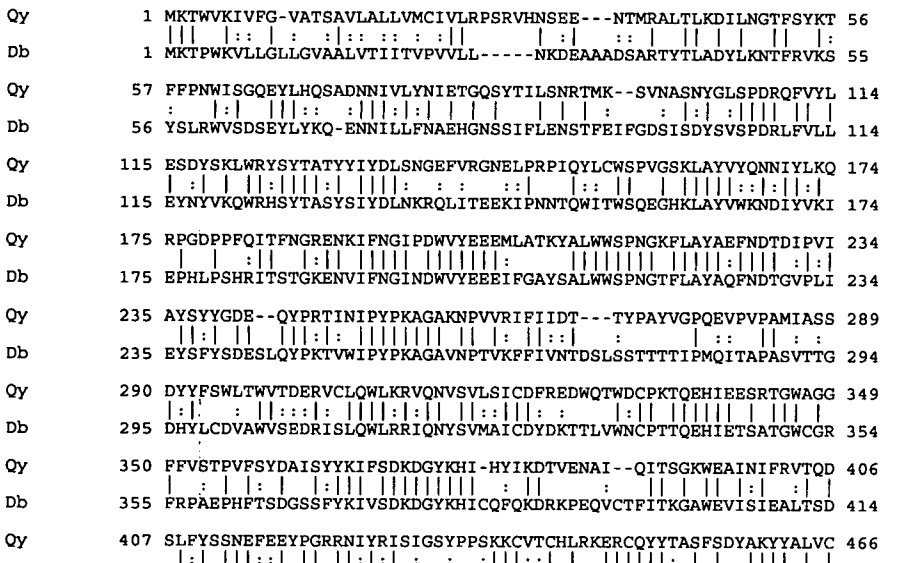
```
Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760


RESULT 2
US-09-265-606-2
; Sequence 2, Application US/09265606
; Patent No. US20020034789A1
;  GENERAL INFORMATION:
;    APPLICANT:  Zimmermann, Rainer; Park, John E.;
;    APPLICANT:  Rettig, Wolfgang; Old, Lloyd J.
;    TITLE OF INVENTION:  ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
;    TITLE OF INVENTION:  ALPHA, AND USES THEREOF
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 2.0 MB  storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/265,606
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,280
;      FILING DATE:  18-MARCH-1996
;      APPLICATION NUMBER:  08/230,491
;      FILING DATE:  20-APRIL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. US20020034789A1man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5330.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  760 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
```

```
US-09-265-606-2

  Query Match             99.6%;  Score 4069;  DB 3;  Length 760;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches  757;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              |||||||||||||||||||||||| |||||||||||||||||| ||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760


RESULT 3
US-10-177-293-136
; Sequence 136, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Glatt, Karen
;  APPLICANT:  Zhao, Xumei
;  APPLICANT:  Gannavarpu, Manjula
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Mertens, Maureen
;  APPLICANT:  Myer, Vic
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Xu, Yongyao
```

```
;  APPLICANT:  Hoersch, Sebastian
;  APPLICANT:  Monahan, John
;  APPLICANT:  Meyers, Rachel E.
;  APPLICANT:  Bast Jr., Robert C.
;  APPLICANT:  Hortobagyi, Gabriel N.
;  APPLICANT:  Pusztai, Lajos
;  APPLICANT:  Meric, Funda
;  APPLICANT:  Sahin, Aysegul
;  APPLICANT:  Mills, Gordon B.
;  TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
;  TITLE OF INVENTION:  PREVENTION, AND THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-038
;  CURRENT APPLICATION NUMBER: US/10/177,293
;  CURRENT FILING DATE:  2002-06-21
;  PRIOR APPLICATION NUMBER: US 60/299,887
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: US 60/301,572
;  PRIOR FILING DATE: 2001-06-27
;  PRIOR APPLICATION NUMBER: US 60/306,501
;  PRIOR FILING DATE: 2001-07-18
;  PRIOR APPLICATION NUMBER: US 60/325,002
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US 60/362,585
;  PRIOR FILING DATE: 2002-03-05
;  PRIOR APPLICATION NUMBER: US 60/xxx,xxx
;  PRIOR FILING DATE: 2002-05-14
;  NUMBER OF SEQ ID NOS: 506
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 136
;   LENGTH: 760
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-177-293-136

  Query Match             99.6%;  Score 4069;  DB 4;  Length 760;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches  757;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              |||||||||||||||||||||||| |||||||||||||||||| ||||||||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760


RESULT 4
US-10-301-822-55
; Sequence 55, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  Berger, Allison
;  APPLICANT:  Guillemette, Tracy L.
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Schlegel, Robert
;  APPLICANT:  Monahan, John E.
;  APPLICANT:  Thibodeau, Stephen N.
;  APPLICANT:  Burgart, Lawrence J.
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
;  TITLE OF INVENTION:  METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF COLON CANCER
;  FILE REFERENCE: MPM01-029P2RNM
;  CURRENT APPLICATION NUMBER: US/10/301,822
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 60/339,971
;  PRIOR FILING DATE: 2001-12-10
;  PRIOR APPLICATION NUMBER: US 60/361,978
;  PRIOR FILING DATE: 2002-03-05
;  PRIOR APPLICATION NUMBER: US 60/381,988
;  PRIOR FILING DATE: 2002-05-20
;  NUMBER OF SEQ ID NOS: 228
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 55
;   LENGTH: 760
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-301-822-55

  Query Match             99.6%;  Score 4069;  DB 4;  Length 760;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches  757;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              |||||||||||||||||||||||| |||||||||||||||||| ||||||||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760


RESULT 5
US-10-884-070A-13
; Sequence 13, Application US/10884070A
; Publication No. US20050170368A1
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi,Avi J.
;  APPLICANT:  Cairns,Belinda
;  APPLICANT:  Dowd,Patrick
;  APPLICANT:  Frantz,Gretchen
;  APPLICANT:  French,Dorothy
;  APPLICANT:  Gonzalez,Lino
;  APPLICANT:  Polakis,Paul
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Wolf,Beni B.
;  APPLICANT:  Wu,Thomas D.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF TUMOR
;  FILE REFERENCE: P5036R1-US
;  CURRENT APPLICATION NUMBER: US/10/884,070A
;  CURRENT FILING DATE:  2004-07-02
;  PRIOR APPLICATION NUMBER: US 60/484,959
;  PRIOR FILING DATE: 2003-07-02
;  NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 13
;   LENGTH: 760
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-884-070A-13

  Query Match             99.6%;  Score 4069;  DB 5;  Length 760;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches  757;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              |||||||||||||||||||||||||| |||||||||||||||||||| ||||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760


RESULT 6
US-10-165-603-4
; Sequence 4, Application US/10165603
; Publication No. US20030021792A1
; GENERAL INFORMATION:
;  APPLICANT: Roben, Paul W.
;  APPLICANT:  Stevens, Anthony C.
;  TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE
;  TITLE OF INVENTION:  PROTEINS
;  FILE REFERENCE: TPTECH.001A
;  CURRENT APPLICATION NUMBER: US/10/165,603
;  CURRENT FILING DATE:  2002-06-07
;  PRIOR APPLICATION NUMBER: 60/297,021
;  PRIOR FILING DATE: 2001-06-08
;  PRIOR APPLICATION NUMBER: 60/305,117
;  PRIOR FILING DATE: 2001-07-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 767
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
US-10-165-603-4
```

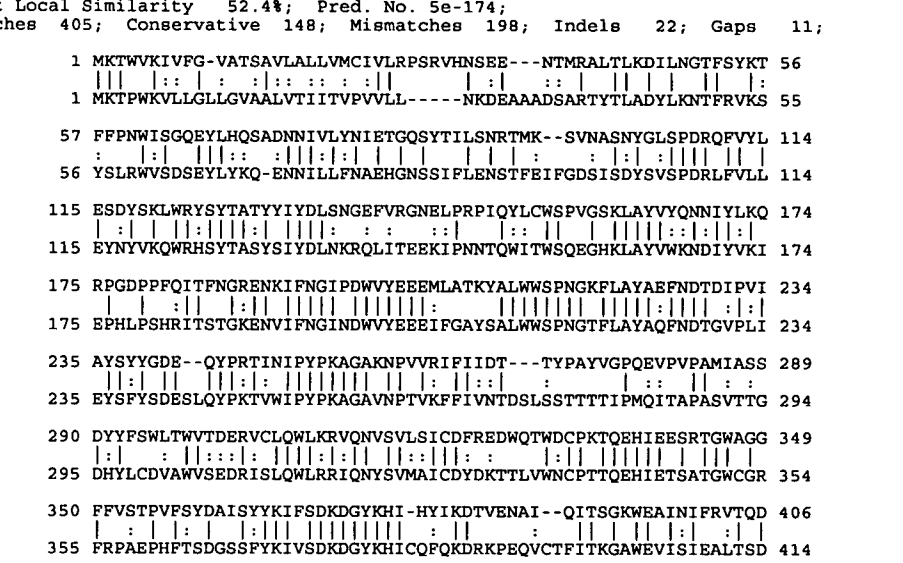
Query Match 53.9%; Score 2204; DB 4; Length 767;
Best Local Similarity 52.4%; Pred. No. 5e-174;
Matches 405; Conservative 148; Mismatches 198; Indels 22; Gaps 11;



RESULT 7
US-10-794-899-38
; Sequence 38, Application US/10794899
; Publication No. US20040146516A1
; GENERAL INFORMATION:
; APPLICANT: Utah Ventures
; TITLE OF INVENTION: Lumen-Exposed Molecules and Methods Targeted Delivery
; FILE REFERENCE: 27110-715
; CURRENT APPLICATION NUMBER: US/10/794,899
; CURRENT FILING DATE: 2004-03-05
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Rattus norvegicus

US-10-794-899-38

Query Match 53.9%; Score 2204; DB 4; Length 767;
Best Local Similarity 52.4%; Pred. No. 5e-174;
Matches 405; Conservative 148; Mismatches 198; Indels 22; Gaps 11;



RESULT 8
US-10-770-712-2
; Sequence 2, Application US/10770712
; Publication No. US20050170333A1
; GENERAL INFORMATION:
; APPLICANT: Vojdani, Aristo
; TITLE OF INVENTION: IDENTIFICATION OF ETIOLOGY OF AUTISM
; FILE REFERENCE: IMSCI2.008A
; CURRENT APPLICATION NUMBER: US/10/770,712
; CURRENT FILING DATE: 2004-02-03
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 767

```
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
US-10-770-712-2

  Query Match             53.8%;  Score 2199;  DB 5;  Length 767;
  Best Local Similarity   52.3%;  Pred. No. 1.3e-173;
  Matches  404;  Conservative  149;  Mismatches  198;  Indels   22;  Gaps   11;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEE---NTMRALTLKDILNGTFSYKT 56
              |||  |:: | :  :|:: :: : :||      | :|   :: |  || | |  ||  |:
Db          1 MKTPWKVLLGLLGVAALVTIITVPVVLL-----NKDEAAADSARTYTLADYLKNTFRVKS 55

Qy         57 FFPNWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMK--SVNASNYGLSPDRQFVYL 114
              :  |:|  |||::  :|||:|:| | | |   | | | :    : |:| :|||| || |
Db         56 YSLRWVSDSEYLYKQ-ENNILLFNAEHGNSSIFLENSTFEIFGDSISDYSVSPDRLFVLL 114

Qy        115 ESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQ 174
              | :| | ||:|||||:| ||||:  : :   ::|   |:: ||  | |||||::|:||:|
Db        115 EYNYVKQWRHSYTASYSIYDLNKRQLITEEKIPNNTQWITWSQEGHKLAYVWKNDIYVKI 174

Qy        175 RPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVI 234
              |  |  :||  |:|| ||||| |||||||:     |||||||| |||||:|||| :|:|
Db        175 EPHLPSHRITSTGKENVIFNGINDWVYEEEIFGAYSALWWSPNGTFLAYAQFNDTGVPLI 234

Qy        235 AYSYYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDT---TYPAYVGPQEVPVPAMIASS 289
              ||:| ||  |||:|: |||||||| || |: ||::|   :       | ::  || : :
Db        235 EYSFYSDESLQYPKTVWIPYPKAGAVNPTVKFFIVNTDSLSSTTTTIPMQITAPASVTTG 294

Qy        290 DYYFSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGG 349
              |:|   : ||:::|: ||||:|:|| ||::|||: :    |:|| |:|||| | ||| |
Db        295 DHYLCDVAWVSEDRISLQWLRRIQNYSVMAICDYDKTTLVWNCPTTREHIETSATGWCGR 354

Qy        350 FFVSTPVFSYDAISYYKIFSDKDGYKHI-HYIKDTVENAI--QITSGKWEAINIFRVTQD 406
              |  : | |: |  |:||| |||||||||  : ||     :   || | || |:|  :| |
Db        355 FRPAEPHFTSDGSSFYKIVSDKDGYKHICQFQKDRKPEQVCTFITKGAWEVISIEALTSD 414

Qy        407 SLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVC 466
              |:| |||::| || ||:|:| :  :  :|||::|  |  |||||: | |  |||| | |
Db        415 YLYYISNEYKEMPGGRNLYKIQLTDH-TNKKCLSCDLNPERCQYYSVSLSKEAKYYQLGC 473

Qy        467 YGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPP 526
              |||:|: |||     :|:::||:|  |:  |:::|:| :::  : ::|   ||:|||||
Db        474 RGPGLPLYTLHRSTDQKELRVLEDNSALDKMLQDVQMPSKKLDFIVLNETRFWYQMILPP 533

Qy        527 QFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLY 586
              ||:|||||||| || |||||   : | :|| :|||| | :::|  ||||: :||||:::
Db        534 HFDKSKKYPLLIDVYAGPCSQKADAAFRLNWATYLASTENIIVASFDGRGSGYQGDKIMH 593

Qy        587 AVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIA 646
              |: ::||  ||||| | |:|::|||:| ||:||||||||||:|: | ||:|:|||||||
Db        594 AINKRLGTLEVEDQIEAARQFLKMGFVDSKRVAIWGWSYGGYVTSMVLGSGSGVFKCGIA 653

Qy        647 VAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVH 706
              ||||| |||| ||||||:|||||  :|||:||:|||||:||| |: |:|||||||||||
Db        654 VAPVSRWEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVH 713

Qy        707 FQNSAQIAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              || ||||:||||:| ||||||||:|::||: |  :  |:|:||:|||:|||||
Db        714 FQQSAQISKALVDAGVDFQAMWYTDEDHGIASSTAHQHIYSHMSHFLQQCFSL 766


RESULT 9
US-10-295-027-590
; Sequence 590, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
;  APPLICANT: Afar, Daniel
;  APPLICANT:  Aziz, Natasha
;  APPLICANT:  Ginsberg, Wendy M.
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Glynne, Richard
;  APPLICANT:  Hevezi, Peter A.
;  APPLICANT:  Mack, David H.
;  APPLICANT:  Murray, Richard
;  APPLICANT:  Watson, Susan R.
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
;  TITLE OF INVENTION:  Methods of Screening for Modulators of Cancer
;  FILE REFERENCE: 018501-012500US
;  CURRENT APPLICATION NUMBER: US/10/295,027
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: US 09/663,733
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/335,394
;  PRIOR FILING DATE: 2001-11-15
;  PRIOR APPLICATION NUMBER: US 60/332,464
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: US 60/334,393
;  PRIOR FILING DATE: 2001-11-29
;  PRIOR APPLICATION NUMBER: US 60/340,376
;  PRIOR FILING DATE: 2001-12-14
;  PRIOR APPLICATION NUMBER: US 60/347,211
;  PRIOR FILING DATE: 2002-01-08
;  PRIOR APPLICATION NUMBER: US 60/347,349
;  PRIOR FILING DATE: 2002-01-10
;  PRIOR APPLICATION NUMBER: US 60/355,250
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: US 60/356,714
;  PRIOR FILING DATE: 2002-02-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1386
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 590
;   LENGTH: 766
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-295-027-590

  Query Match             53.6%;  Score 2191.5;  DB 4;  Length 766;
  Best Local Similarity   51.6%;  Pred. No. 5.4e-173;
  Matches  396;  Conservative  148;  Mismatches  212;  Indels   11;  Gaps    7;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  ||: | :  :|:: :: : :||       ::  :: :  || | |  |:  | :
Db          1 MKTPWKILLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
              |||  |||::  :|||:::| | | |   | | |       |  :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
              | ::||: |:|: |:||| |||||||: :   ||||||| |||||:||||::|:| ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :|||||| || |: |:::|   : |      ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
              :|| | ||: ||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||::  || ||:|:|: :  |      |::| |  ||||||: |||  |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478
```

```
Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              :|||    |: :::||:|  |:  |:|:|:| :::  : ::|   ||:||||| ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||| | |:| :|||:| |||||||||||||||:|: | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| ||||||:||||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |:||||:  ||||||||:|::||: |  :  |:||||:||:||||||
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765
```

RESULT 10
US-09-265-606-3
; Sequence 3, Application US/09265606
; Patent No. US20020034789A1
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20020034789A1man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-265-606-3

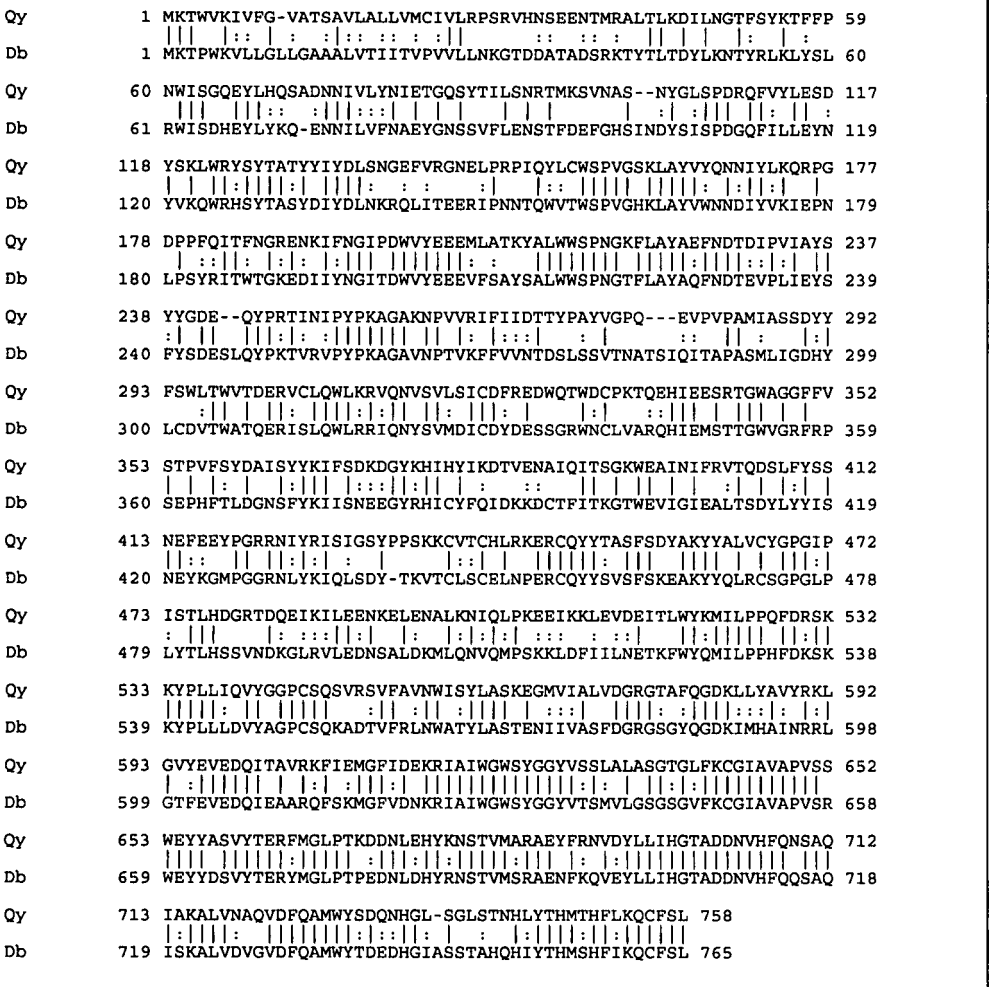
Query Match 53.6%; Score 2190.5; DB 3; Length 766;
Best Local Similarity 51.5%; Pred. No. 6.6e-173;
Matches 395; Conservative 149; Mismatches 212; Indels 11; Gaps 7;

```
Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  |:: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db          1 MKTPWKVLLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60
```

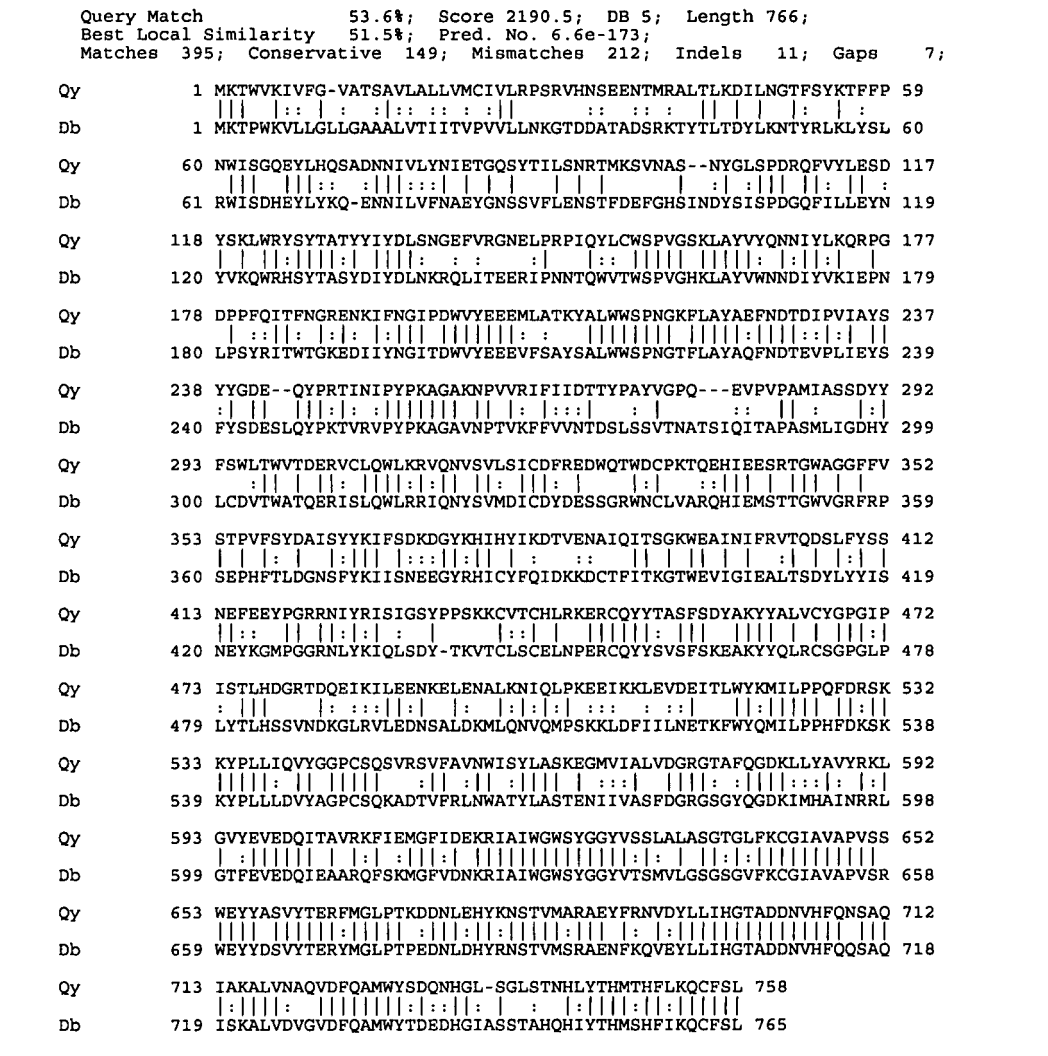
```
Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
              |||  |||::  :|||:::| | | |   | | |       | :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
              | ::||: |:|: |:||| ||||||: :   |||||||| |||||:||||::|:| ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :|||||| || |: |:::|   : |      ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
              :|| | ||: ||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||::  || ||:|:| :  |      |::| |  ||||||: |||  |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||    |: :::||:|  |:  |:|:|:| :::  : ::|   ||:||||| ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||| | |:| :|||:| |||||||||||||||:|: | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| ||||||:||||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |:||||:  ||||||||:|::||: |  :  |:||||:||:||||||
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765
```

RESULT 11
US-09-993-959-1
; Sequence 1, Application US/09993959
; Publication No. US20030165489A1
; GENERAL INFORMATION:
; APPLICANT: Grouzmann, Eric
; APPLICANT: Lacroix, Jean-Silvain
; APPLICANT: Monod, Michel
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
; FILE REFERENCE: 81985/276823
; CURRENT APPLICATION NUMBER: US/09/993,959
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-993-959-1

Query Match 53.6%; Score 2190.5; DB 3; Length 766;
Best Local Similarity 51.5%; Pred. No. 6.6e-173;
Matches 395; Conservative 149; Mismatches 212; Indels 11; Gaps 7;



RESULT 12
US-10-952-459-18
; Sequence 18, Application US/10952459
; Publication No. US20050074805A1
; GENERAL INFORMATION:
; APPLICANT: Kochan, Jarema Peter
; APPLICANT: Martin, Mitchell Lee
; APPLICANT: Rosinski, James Andrew
; TITLE OF INVENTION: Specific Markers for Diabetes
; FILE REFERENCE: 21270US1
; CURRENT APPLICATION NUMBER: US/10/952,459
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-952-459-18



RESULT 13
US-10-631-467-565
; Sequence 565, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genox Research Inc.
; TITLE OF INVENTION: Method for testing for broncheal asthma, or chronic obstructive p
; TITLE OF INVENTION: disease
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631,467
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-229312
; PRIOR FILING DATE: 2002-08-06

; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 565
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-467-565

Query Match 53.6%; Score 2190.5; DB 5; Length 766;
Best Local Similarity 51.5%; Pred. No. 6.6e-173;
Matches 395; Conservative 149; Mismatches 212; Indels 11; Gaps 7;

```
Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  |:: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db          1 MKTPWKVLLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
              |||  |||::  :|||:::| | | |    | | |       |  :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | |:||:|||:| |||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
              | ::||: |:|: |:||| |||||||: :   |||||||| |||||:||||::|:| ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :|||||| || |: |:::|   : |      ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
              :|| | ||: ||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||::  || ||:|:| :  |      |::| |  ||||||: |||  |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||    |: :::||:|  |:  |:|:|:| :::  : ::|   ||:|||||  ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||| | |:| :|||:| |||||||||||||:|: | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| ||||||:||||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |:||||:  |||||||:|::||: |  :  |:||||:||:||||||
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765
```

RESULT 14
US-10-002-593-6
; Sequence 6, Application US/10002593
; Publication No. US20020137120A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTIN

; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Atty Docket No. US20020137120A1 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/002,593
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-593-6

Query Match 53.5%; Score 2187.5; DB 4; Length 766;
Best Local Similarity 51.6%; Pred. No. 1.2e-172;
Matches 396; Conservative 148; Mismatches 212; Indels 11; Gaps 7;

```
Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  ||: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db          1 MKTPWKILLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
              |||  |||::  :|||:::| | | |    | | |       |  :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
              | ::||: |:|: |:||| |||||||: :   |||||||| |||||:||||::|:| ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :|||||| || |: |:::|   : |      ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
              :|| | ||: ||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||::  || ||:|:| :  |      |::| |  ||||||: |||  |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLIDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||    |: :::||:|  |:  |:|:|:| :::  : ::|   ||:|||||  ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||| | |:| :|||:| |||||||||||||:|: | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| ||||||:||||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |:||||:  |||||||:|::||: |  :  |:||||:||:||||||
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765
```

RESULT 15
US-10-165-603-7

```
; Sequence 7, Application US/10165603
; Publication No. US20030021792A1
; GENERAL INFORMATION:
;  APPLICANT: Roben, Paul W.
;  APPLICANT:  Stevens, Anthony C.
;  TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE
;  TITLE OF INVENTION:  PROTEINS
;  FILE REFERENCE: TPTECH.001A
;  CURRENT APPLICATION NUMBER: US/10/165,603
;  CURRENT FILING DATE:  2002-06-07
;  PRIOR APPLICATION NUMBER: 60/297,021
;  PRIOR FILING DATE: 2001-06-08
;  PRIOR APPLICATION NUMBER: 60/305,117
;  PRIOR FILING DATE: 2001-07-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 766
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-165-603-7

  Query Match             53.5%;  Score 2187.5;  DB 4;  Length 766;
  Best Local Similarity   51.6%;  Pred. No. 1.2e-172;
  Matches  396;  Conservative  148;  Mismatches  212;  Indels   11;  Gaps    7;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  ||: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db          1 MKTPWKILLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
               |||  |||::  :|||:::| | | |   | | |      |  :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
               | ::||: |:|: |:||| |||||||: :   ||||||| |||||:||||::|:| ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :|||||| || |: |:::|   : |      ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
                :|| | ||: ||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||::  || ||:|:| :  |     |::| |  |||||: |||  |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLIDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||    |: :::||:|  |:  |:|:|:| :::  : ::|   ||:|||| ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||  | |:| :|||:| ||||||||||||||:|: | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLLIHGTADDNVHFQNSAQ 712
              |||| ||||||:||||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |:||||:  |||||||||:|::||: |  :  |:||||:||:||||||
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765


Search completed: January 28, 2006, 00:46:15
Job time : 118 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2006, 00:35:22 ; Search time 34 Seconds
(without alignments)
242.062 Million cell updates/sec

Title: US-10-681-659-1
Perfect score: 4086
Sequence: 1 MKTWVKIVFGVATSAVLALL.....TNHLYTHMTHFLKQCFSLSD 760

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4069	99.6	760	7	US-11-186-284-55	Sequence 55, Appl
2	2190.5	53.6	766	6	US-10-522-789-2	Sequence 2, Appli
3	2165.5	53.0	762	7	US-11-116-939-13	Sequence 13, Appl
4	456	11.2	882	7	US-11-151-601-20	Sequence 20, Appl
5	257.5	6.3	657	7	US-11-179-977-1	Sequence 1, Appli
6	119.5	2.9	2710	7	US-11-051-453-41	Sequence 41, Appl
7	112.5	2.8	668	6	US-10-454-437-118	Sequence 118, App
8	111.5	2.7	2725	7	US-11-113-424-52	Sequence 52, Appl
9	108	2.6	1436	6	US-10-453-372-1094	Sequence 1094, Ap
10	108	2.6	1593	6	US-10-453-372-1092	Sequence 1092, Ap
11	105.5	2.6	710	7	US-11-151-601-23	Sequence 23, Appl
12	104.5	2.6	433	7	US-11-110-851-65	Sequence 65, Appl
13	103	2.5	2671	6	US-10-876-787-6	Sequence 6, Appli
14	101.5	2.5	621	6	US-10-467-657-1554	Sequence 1554, Ap
15	100	2.4	282	6	US-10-878-556A-125	Sequence 125, App
16	98.5	2.4	2367	7	US-11-051-453-42	Sequence 42, Appl
17	97	2.4	256	7	US-11-179-977-2	Sequence 2, Appli
18	96.5	2.4	527	7	US-11-052-554A-304	Sequence 304, App
19	96.5	2.4	750	7	US-11-070-627-4	Sequence 4, Appli
20	96	2.3	72	7	US-11-151-601-22	Sequence 22, Appl
21	94	2.3	1206	6	US-10-467-657-72	Sequence 72, Appl
22	94	2.3	1206	6	US-10-467-657-3892	Sequence 3892, Ap
23	92	2.3	728	6	US-10-467-657-8376	Sequence 8376, Ap
24	91.5	2.2	1647	7	US-11-052-554A-260	Sequence 260, App
25	91	2.2	1219	7	US-11-054-281-28	Sequence 28, Appl
26	91	2.2	1219	7	US-11-054-281-106	Sequence 106, App
27	90	2.2	867	6	US-10-454-437-104	Sequence 104, App
28	90	2.2	3063	7	US-11-186-284-26	Sequence 26, Appl
29	89.5	2.2	468	6	US-10-467-657-3478	Sequence 3478, Ap
30	89.5	2.2	545	6	US-10-467-657-1408	Sequence 1408, Ap
31	89	2.2	754	6	US-10-793-626-1296	Sequence 1296, Ap
32	89	2.2	862	7	US-11-054-281-109	Sequence 109, App
33	88.5	2.2	527	6	US-10-821-234-1574	Sequence 1574, Ap
34	88.5	2.2	841	7	US-11-052-554A-102	Sequence 102, App
35	88.5	2.2	1045	7	US-11-113-424-54	Sequence 54, Appl
36	88.5	2.2	1094	6	US-10-821-234-1097	Sequence 1097, Ap
37	88.5	2.2	2376	7	US-11-096-051-4	Sequence 4, Appli
38	88.5	2.2	2715	7	US-11-096-051-2	Sequence 2, Appli
39	88.5	2.2	2721	7	US-11-096-051-10	Sequence 10, Appl
40	88.5	2.2	2725	7	US-11-096-051-8	Sequence 8, Appli
41	88	2.2	1104	7	US-11-099-855-11	Sequence 11, Appl
42	87.5	2.1	690	7	US-11-052-554A-232	Sequence 232, App
43	87.5	2.1	721	7	US-11-052-554A-253	Sequence 253, App
44	87.5	2.1	875	6	US-10-859-643-743	Sequence 743, App
45	87.5	2.1	875	6	US-10-859-643-745	Sequence 745, App

ALIGNMENTS

RESULT 1
US-11-186-284-55
; Sequence 55, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-186-284-55

Query Match 99.6%; Score 4069; DB 7; Length 760;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 757; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
```

```
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
```

RESULT 2
US-10-522-789-2
; Sequence 2, Application US/10522789
; Publication No. US20050260732A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO., LTD.
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522,789
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-522-789-2

Query Match 53.6%; Score 2190.5; DB 6; Length 766;
Best Local Similarity 51.5%; Pred. No. 1.4e-175;
Matches 395; Conservative 149; Mismatches 212; Indels 11; Gaps 7;

```
Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
Db          1 MKTPWKVLLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
```

```
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
Db        420 NEYKGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765
```

RESULT 3
US-11-116-939-13
; Sequence 13, Application US/11116939
; Publication No. US20050265995A1
; GENERAL INFORMATION:
; APPLICANT: Stephen Tomlinson
; APPLICANT: Richard J. Quigg
; TITLE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS
; FILE REFERENCE: 19113.0115U2
; CURRENT APPLICATION NUMBER: US/11/116,939
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: 60/565,907
; PRIOR FILING DATE: 2004-04-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 762
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; note=synthetic
; OTHER INFORMATION: construct
US-11-116-939-13

Query Match 53.0%; Score 2165.5; DB 7; Length 762;

```
  Best Local Similarity   51.2%;  Pred. No. 1.7e-173;
  Matches  394;  Conservative  148;  Mismatches  208;  Indels   19;  Gaps    9;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  |:: | :  :|:: :: : :||       ::  :: :  || | |       | :
Db          1 MKTPWKVLLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYL------KNYLK 54

Qy         60 --NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLE 115
                |||  |||::  :|||:::| | | |   | | |       |  :| :||| ||: ||
Db         55 KLRWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLE 113

Qy        116 SDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQR 175
               :| | ||:|||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|
Db        114 YNYVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIE 173

Qy        176 PGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIA 235
              |  | ::||: |:|: |:||| ||||||: :   |||||||| |||||:||||::|:|
Db        174 PNLPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIE 233

Qy        236 YSYYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSD 290
              ||:| ||  |||:|: :|||||| || |: |:::|   : |       ::  || :   |
Db        234 YSFYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGD 293

Qy        291 YYFSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGF 350
               :|   :|| | ||: ||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        294 HYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRF 353

Qy        351 FVSTPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFY 410
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:|
Db        354 RPSEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYY 413

Qy        411 SSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPG 470
              |||::  || ||:|:| :  |      |::| |  ||||||: |||  |||| | | |||
Db        414 ISNEYKGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPG 472

Qy        471 IPISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDR 530
               :|: |||     |: :::||:|  |:  |:|:|:|: :::  : ::|   ||:|||| ||:
Db        473 LPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDK 532

Qy        531 SKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYR 590
              |||||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |
Db        533 SKKYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINR 592

Qy        591 KLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPV 650
              :|| :|||||| | |:|  :|||:| ||||||||||||||:|: | ||:|:|||||||||
Db        593 RLGTPEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPV 652

Qy        651 SSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNS 710
              | |||| ||||||:|||||  :|||:||:|||||:||| |: |:|||||||||||||| |
Db        653 SRWEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQS 712

Qy        711 AQIAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |||:||||:  |||||||:|::||: |   :  |:||||:||:|||||
Db        713 AQISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 761


RESULT 4
US-11-151-601-20
; Sequence 20, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  Meyers, Rachel E.
;  APPLICANT:  Olandt,  Peter J.
;  APPLICANT:  Kapeller-Libermann, Rosana
;  APPLICANT:  Curtis, Rory A. J.
;  APPLICANT:  Williamson, Mark
;  APPLICANT:  Weich, Nadine
;  TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
;  TITLE OF INVENTION:  AND PROTEASE FAMILY MEMBERS AND USES THEREOF
;  FILE REFERENCE: MPI00-054P1RCP1OMNIDV1M
;  CURRENT APPLICATION NUMBER: US/11/151,601
;  CURRENT FILING DATE:  2005-06-13
;  PRIOR APPLICATION NUMBER: US 10/170,789
;  PRIOR FILING DATE: 2002-06-13
;  PRIOR APPLICATION NUMBER: US 09/797,039
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: PCT/US01/06525
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: US 60/186,061
;  PRIOR FILING DATE: 2000-02-29
;  PRIOR APPLICATION NUMBER: US 09/882,166
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: PCT/US01/19269
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: US 60/212,078
;  PRIOR FILING DATE: 2000-06-15
;  PRIOR APPLICATION NUMBER: US 09/934,406
;  PRIOR FILING DATE: 2001-08-21
;  PRIOR APPLICATION NUMBER: PCT/US01/26052
;  PRIOR FILING DATE: 2001-08-21
;  PRIOR APPLICATION NUMBER: US 60/226,740
;  PRIOR FILING DATE: 2000-08-21
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 45
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
;   LENGTH: 882
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-151-601-20

  Query Match             11.2%;  Score 456;  DB 7;  Length 882;
  Best Local Similarity   24.1%;  Pred. No. 3.9e-30;
  Matches  169;  Conservative  114;  Mismatches  268;  Indels  150;  Gaps   27;

Qy        162 LAYVYQNNIYLKQRPGDPPFQITF------NGRENKIFNGIPDWVYEEEMLATKYA-LWW 214
              :|::: |:|::          ::|:      |  |:    |: :| :||    :|:  ||
Db        212 IAFIHSNDIWISNIVTREERRLTYVHNELANMEEDARSAGVATFVLQEEF--DRYSGYWW 269

Qy        215 SP-------NGKFL--AYAEFNDTDIPVIAYSYYGDEQYPRTINIPYPKAGAKNPVVRI- 264
              |        || |   | | ::::: :|  :    |   |  :  ||| |  || |
Db        270 CPKAETTPSGGKILRILYEENDESEVEIIHVTSPMLET-RRADSFRYPKTGTANPKVTFK 328

Qy        265 ----------FIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVTDERVCLQWL----- 309
                       |||      : | |:     :     | :   |  : :     |
Db        329 MSEIMIDAEGRIIDVIDKELIQPFEI-----LFEGVEYIARAGWTPEGKYAWSILLDRSQ 383

Qy        310 KRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYDA-----ISY 364
              |:| | :         :   :  :  | : :|           ||:  |:      |:
Db        384 TRLQIVLISPELFIPVEDDVMERQRLIESVPDS----------VTPLIIYEETTDIWINI 433

Qy        365 YKIF-----------------SDKDGYKHIHYI---------------------KDTVE 385
              : ||                   | |::|:: |                       |  ::
Db        434 HDIFHVFPQSHEEEIEFIFASECKTGFRHLYKITSILKESKYKRSSGGLPAPSDFKCPIK 493

Qy        386 NAIQITSGKWEAI-----NI----------FRVTQDSLFYSSNEFEEYPGRRNIYRISIG 430
              | ||||:|| :     ||          |  |:||        |     : |::
Db        494 EEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHHLYVVSYVNPGEVTRLTDR 553

Qy        431 SYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGRTDQE----IK 486
              |  |  |::  |     | :: : :|:    : :  |     :|:  |  | :
Db        554 GYSHS-CCISQH-----CDFFISKYSNQKNPHCVSLY----KLSSPEDDPTCKTKEFWAT 603

Qy        487 ILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQVYGGPCS 546
              ||:    | :     | |          ||:  :  |      |||| :: :||||
Db        604 ILDSAGPLPD-----YTPPEIFSFESTTGFTLYGMLYKPHDLQPGKKYPTVLFIYGGPQV 658

Qy        547 QSVRS------VFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              . | | :       | :| :: |    | |: ::| ||:  :| |   |    |:|  |::||
Db        659 QLVNNRFKGVKYFRLNTLASL----GYVVVVIDNRGSCHRGLKFEGAFKYKMGQIEIDDQ 714

Qy        601 ITAVRKFI-EMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASV 659
              :  ::       |||  |: | ||||||:| :||   : :|:  || |||: | :| :
Db        715 VEGLQYLASRYDFIDLDRVGIHGWSYGGYLSLMALMQRSDIFRVAIAGAPVTLWIFYDTG 774
```

```
Qy        660 YTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVD--YLLIHGTADDNVHFQNSAQIAKAL 717
Db        775 YTERYMGHP--DQNEQGYYLGSVAMQAEKFPSEPNRLLLLHGFLDENVHFAHTSILLSFL 832

Qy        718 VNAQVDFQAMWYSDQNHGL----SGLSTNHLYTHMTHFLKQ 754
Db        833 VRAGKPYDLQIYPQERHSIRVPESG---EHYELHLLHYLQE 870


RESULT 5
US-11-179-977-1
; Sequence 1, Application US/11179977
; Publication No. US20050249789A1
; GENERAL INFORMATION:
;  APPLICANT: Genencor International, Inc.
;  TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
;  FILE REFERENCE: GC511-PCT
;  CURRENT APPLICATION NUMBER: US/11/179,977
;  CURRENT FILING DATE:  2005-07-12
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 657
;   TYPE: PRT
;   ORGANISM: Bacillus
US-11-179-977-1

  Query Match             6.3%;  Score 257.5;  DB 7;  Length 657;
  Best Local Similarity   21.2%;  Pred. No. 1.1e-13;
  Matches  153;  Conservative  121;  Mismatches  276;  Indels  171;  Gaps   31;

Qy         71 SADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSKLWRYSYTATY 130
Db         41 SYTSNIWIYETKTGGSVP----WTHGEKRSTDPRWSPDGRTLAFISD-----REGDAAQL 91

Qy        131 YIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQ----NNIYLKQRPGDPPFQ-ITF 185
Db         92 YIMSTEGGEARKLTDIPYGVSKPLWSPDGESILVTISLGEGESIDDREKTEQDSYEPVEV 151

Qy        186 NG----RENKIFNGIPDWVYEEEMLATKYA---------------LWWSPNGKFLAY-AE 225
Db        152 QGLSYKRDGK---GLTRGAYAQLVLVSVKSGEMKELTSHKADHGDPAFSPDGKWLVFSAN 208

Qy        226 FNDTDIPVIAYSYYGDEQYPRTINIPYPKAGAKNPVV--------RIFIIDTTYPAYVGP 277
Db        209 LTETD----------DASKPHDVYIMSLESGDLKQVTPHRGSFGSSSFSPDGRYLALLGN 258

Qy        278 QEVPVPAMIASSDYYFSWLTWVTDERV-CLQWLKRVQNVSVLSICDFREDWQTWDCPKTQ 336
Db        259 EKEYKNATLSK-----AWLYDIEQGRLTCLTEMLDVHLADAL------------------ 295

Qy        337 EHIEESRTGWAGGFFVSTPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWE 396
Db        296 --IGDSLIGGA----EQRPIWTKDSQGFYVIGTDQ-GSTGIYYI--SIEGLVYPIRLEKE 346

Qy        397 AINIFRVTQDSLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFS 456
Db        347 YINSFSLSPDE-----------------------------------------QHFIASVT 365

Qy        457 DYAKYYALVCYGPGIPISTLHDGRTDQEIKILEENKELENA------LKNIQLPKEEIKK 510
Db        366 K-----------PDRP-SELY------SIPLGQEEKQLTGANDKFVREHTISIP-EEIQY 406

Qy        511 LEVDEITLWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIA 570
Db        407 ATEDGVMVNGWLMRPAQMEGETTYPLILNIHGGPHMMYGHTYF--HEFQVLAAK-GYAVV 463

Qy        571 LVDGRGTAFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEMG-FIDEKRIAIWGWSYGGYV 629
Db        464 YINPRGSHGYGQEFVNAVRGDYGGKDYDDVMQAVDEAIKRDPHIDPKRLGVTGGSYGGFM 523

Qy        630 SSLALASGTGLFKCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEH--YKNSTVM---A 684
```

```
Db        524 TNW-IVGQTNRFKAAVTQRSISNWISFHGVSD---IGYFFTDWQLEHDMFEDTEKLWDRS 579

Qy        685 RAEYFRNVD--YLLIHGTADDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLSTN 742
Db        580 PLKYAANVETPLLILHGERDDRCPIEQAEQLFIALKKMGKETKLVRFPNASHNLS--RTG 637

Qy        743 H 743
Db        638 H 638


RESULT 6
US-11-051-453-41
; Sequence 41, Application US/11051453
; Publication No. US20050287150A1
; GENERAL INFORMATION:
;  APPLICANT: AMBROSINO, DONNA
;  APPLICANT:  BABCOCK, GREGORY J.
;  APPLICANT:  BROERING, THERESA
;  APPLICANT:  GRAZIANO, ROBERT
;  APPLICANT:  HERNANDEZ, HECTOR JAVIER
;  APPLICANT:  LOWY, ISRAEL
;  APPLICANT:  MANDELL, ROBERT
;  APPLICANT:  MOLRINE, DEBORAH
;  APPLICANT:  THOMAS, JR., WILLIAM D.
;  APPLICANT:  ZHANG, HUI-FEN
;  TITLE OF INVENTION: ANTIBODIES AGAINST CLOSTRIDIUM DIFFICILE TOXINS AND
;  TITLE OF INVENTION:  USES THEREOF
;  FILE REFERENCE: MJI-001
;  CURRENT APPLICATION NUMBER: US/11/051,453
;  CURRENT FILING DATE:  2005-02-04
;  PRIOR APPLICATION NUMBER: 60/542,357
;  PRIOR FILING DATE: 2004-02-06
;  PRIOR APPLICATION NUMBER: 60/613,854
;  PRIOR FILING DATE: 2004-09-28
;  NUMBER OF SEQ ID NOS: 82
;  SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 41
;   LENGTH: 2710
;   TYPE: PRT
;   ORGANISM: Clostridium difficile
US-11-051-453-41

  Query Match             2.9%;  Score 119.5;  DB 7;  Length 2710;
  Best Local Similarity   18.7%;  Pred. No. 0.31;
  Matches  145;  Conservative   88;  Mismatches  273;  Indels  271;  Gaps   35;

Qy         37 NTMRALTLKDILNGTFSY-----------------KTFFP-----NWISGQEYLHQSA- 72
Db       1867 NTGAALTSYKIINGKHFYFNNDGVMQLGVFKGPDGFEYFAPANTQNNNIEGQAIVYQSKF 1926

Qy         73 ----------DNN---IVLYNIETGQSYTILSNRTMKSV-----NASNYGLSPDRQFVYL 114
Db       1927 LTLNGKKYYFDNNSKAVTGWRIINNEKYYFNPNNAIAAVGLQVIDNNKYYFNPDTAII-- 1984

Qy        115 ESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQ 174
Db       1985 ----SKGWQTVNGSRYYF---------------------------------------- 1998

Qy        175 RPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNG-KFLAYAEFNDTDIPV 233
Db       1999 ---DTDTAIAFNG--YKTIDG--KHFYFDSDCVVKIGVFSTSNGFEYFAPANTYNNNIEG 2051

Qy        234 IAYSYYGDEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYF 293
Db       2052 QAIVY---QSKFLTLN------GKK------YYFDNNSKAVTGWQ------TIDSKKYYF 2090

Qy        294 SWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPK--TQEHIEESRTGWAGGFF 351
Db       2091 NTNT----------------------AEAATGWQTIDGKKYYFNTNTAEAATGWQ---- 2123

Qy        352 VSTPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYS 411
```

```
Db       2124 ------TIDGKKYY--FNTNTAIASTGY---------TIINGKHFYFNTDGIMQIGVFKG 2166

Qy        412 SNEFEEY-PGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFS---------DYAKY 461

Db       2167 PNGFEYFAPANTDANNIEGQAILYQNEFLTLNGKK----YYFGSDSKAVTGWRIINNKKY 2222

Qy        462 YALVCYGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYK 521

Db       2223 Y----FNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITI------------ERNNFY- 2264

Qy        522 MILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQG 581

Db       2265 ------FDANNESKMVTGVFKGP----------NGFEYFAP--------ANTHNNNIEG 2299

Qy        582 DKLLYAVYRKLGVYEVEDQITAVRKFIEMG----FIDEKRIAIWGW------SYGGYVSS 631

Db       2300 QAIVYQ----------------NKFLTLNGKKYYFDNDSKAVTGWQTIDGKKYYFNLNT 2342

Qy        632 LALASGTGLFKCGIAVAPVSSWEYYASVYT-ERFMGLPTKDDNLEHYKNSTVMARAEY-- 688

Db       2343 AEAATGW---------QTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTS 2393

Qy        689 ------FRNVDYLLIHGTADDNVHFQNSAQIAKALVNA-QVDFQAMWYSDQNHGLSG 738

Db       2394 INGKHFYFNTDGIMQIGVFKGPNGFE---YFAPANTDANNIEGQAILYQNKFLTLNG 2447
```

RESULT 7
US-10-454-437-118
; Sequence 118, Application US/10454437
; Publication No. US20050277115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CPCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 118
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-118

Query Match 2.8%; Score 112.5; DB 6; Length 668;
Best Local Similarity 21.4%; Pred. No. 0.16;
Matches 77; Conservative 43; Mismatches 127; Indels 113; Gaps 16;

```
Qy        476 LHDGRTDQEIKILEENKELENALKNIQLPKE------EIKKLEVDEITLWYKMI------ 523

Db        337 LNDPTTEHE--------------HIDLPEHVTAHVVATSPLDGDEI--WVQAASFTEAP 379

Qy        524 ------LPPQFDRSKKYPL-----------------------------------LIQV 540

Db        380 TLLRAELPGALEAVKKAPLQFENAGQETRQHWATSADGTKIPYFITGAFEEEPQNTLVHA 439

Qy        541 YGG------PCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGV 594

Db        440 YGGFEVSLTPSHSPTR---GIAWL-----EKGYYFVEANLRGGGEFGPE-WHSQATKLNR 490

Qy        595 YEV-EDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVS-- 651

Db        491 MKVWEDHRAVLADLVERGYATPEQIAIRGGSNGGLLTSGALTQYPEAFGAAVVQVPLADM 550

Qy        652 ----SWEYYASVYTERFMGLPTKDDN---LEHYKNSTVMARAEYFRNVDYLLIHGTADDN 704

Db        551 LRYHTWSAGASWMAE--YGNPDDPEERAVIEQYSPVQAVVGVEKRIYPPALVTTSTRDDR 608

Qy        705 VHFQNSAQIAKALVNA--QVDFQAMWYSDQNHGLSGLSTNH--------LYTHMTHFLKQ 754

Db        609 VHPAHARLFAQALLDAGQAVDY----YENTEGGHAGAADNKQTAFVESLIYTWIEKTLDQ 664
```

RESULT 8
US-11-113-424-52
; Sequence 52, Application US/11113424
; Publication No. US20050260713A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/11/113,424
; CURRENT FILING DATE: 2005-04-21
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
; LENGTH: 2725
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-113-424-52

Query Match 2.7%; Score 111.5; DB 7; Length 2725;
Best Local Similarity 19.3%; Pred. No. 1.5;
Matches 146; Conservative 77; Mismatches 237; Indels 297; Gaps 38;

```
Qy        103 YGLSPDRQFVYLESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVG--S 160

Db       2151 YEYDADGQLQTVSVNDKTQWRYSYDLNGNINLLSHGKSAR----LTPLRYDLRDRITRLG 2206

Qy        161 KLAYVYQNNIYLKQRPGDPPFQITFNGRENKIFNGIPDWVYE------EEMLATKYALWW 214

Db       2207 EIQYKMDEDGFLRQR-GNDIFEYNSNGLLQKAYNKASGWTVQYYYDGLGRRVASKSSL-- 2263
```

```
Qy        215 SPNGKFLA--YAE----------FNDTDIPVIAYSY-----------YGDEQYPRTINI 250
Db       2264 ---GQHLQFFYADLTNPIRVTHLYNHTSSEITSLYYDLQGHLIAMELSSGEEYYVACDNT 2320

Qy        251 PYPKA--GAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVTDERVCLQW 308
Db       2321 GTPLAVFSSRGQVIK----EILYTPY--------------GDIYH--------------- 2347

Qy        309 LKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYDAISYYKIF 368
Db       2348 ------------DTYPDFQV-------------IIGFHGGL--------YDFLTKLVHL 2373

Qy        369 SDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPGRRNIYRIS 428
Db       2374 GQRD---------------YDVVAGRWTTPNHHIWKQLNLL---------PKPFNLYSFE 2409

Qy        429 IGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYA-----LVCYG-------PGIPISTL 476
Db       2410 -NNYPVGK-----------------IQDVAKYTTDIRSWLELFGFQLHNVLPGFPKPEL 2450

Qy        477 HDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLW--YKMILPPQFDRSK-- 532
Db       2451 ------------ENLELTYELLRLQTKTQE----------WDPGKTILGIQCELQKQL 2486

Qy        533 -------KYPLLIQVYGGPCSQ--------SVRSVFAVNWISYLASKEGMVIALVDGRGT 577
Db       2487 RNFISLDQLPMTPRYNDGRCLEGGKQPRFAAVPSVFG-KGIKF-AIKDGIVTA--DIIGV 2542

Qy        578 AFQGDKLLYAVYRKLGVYE-----VEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSL 632
Db       2543 ANEDSRRLAAILNNAHYLENLHFTIEGRDT--HYFIKLGSLEE--------------DL 2585

Qy        633 ALASGTG---LFKCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYF 689
Db       2586 VLIGNTGGRRILENGVNVT-VSQMTSVLNGRTRRF------------------------ 2619

Qy        690 RNVDYLLIHGTADDNVHFQNSA--------QIAKALVNAQVDFQAMWYSDQNHGLSG--- 738
Db       2620 --ADIQLQHGALCFNIRYGTTVEEEKNHVLEIARQRAVAQA-----WTKEQRRLQEGEEG 2672

Qy        739 ------------LSTNHLYTHMTHF---LKQCFSLSD 760
Db       2673 IRAWTEGEKQQLLSTGRVQGYDGYFVLSVEQYLELSD 2709


RESULT 9
US-10-453-372-1094
; Sequence 1094, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
;  APPLICANT: Alsobrook, et al.
;  TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
;  FILE REFERENCE: 21402-589 A
;  CURRENT APPLICATION NUMBER: US/10/453,372
;  CURRENT FILING DATE:  2003-06-03
;  PRIOR APPLICATION NUMBER: 09/789390
;  PRIOR FILING DATE: 2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 09/823187
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/195792
;  PRIOR FILING DATE: 2000-03-10
;  PRIOR APPLICATION NUMBER: 09/839446
;  PRIOR FILING DATE: 2001-03-19
;  PRIOR APPLICATION NUMBER: 60/199476
;  PRIOR FILING DATE: 2000-03-25
;  PRIOR APPLICATION NUMBER: 09/863776
;  PRIOR FILING DATE: 2001-05-23
;  PRIOR APPLICATION NUMBER: 60/208263
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: 09/939398
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: 60/227800
;  PRIOR FILING DATE: 2000-08-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1609
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 1094
;   LENGTH: 1436
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-453-372-1094

  Query Match             2.6%;  Score 108;  DB 6;  Length 1436;
  Best Local Similarity   19.0%;  Pred. No. 1.1;
  Matches  148;  Conservative  109;  Mismatches  285;  Indels  236;  Gaps   41;

Qy         96 KSVNASNYGLSPDR-----------QFVYLESDYSKLWRYSYTATYYIYDLSNGEFVRG 143
Db         52 KSLPGSNIDLQVSAASNSLCALWAVDQSVLLLRNYGQLSAQTVYSQLYSREL-HGYYFRG 110

Qy        144 NELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRP-----GDPPFQITFNGRENKIFNGIPD 198
Db        111 LNLEDGLKVPCLED-----EHILYNGIY--YTPAWADFGKDGYDLVKDPQNNRIFQ---- 159

Qy        199 WVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYGDEQYPRTINIPYPKAGAK 258
Db        160 -----------------RQNVTSFRNITQLSFQLISEPMFGD--YWIVV-----KRNSR 194

Qy        259 NPVVRIFIIDTTYPAYVGPQ-EVPV--PAMIASSDYYFSWLTWVTDERVCLQWL--KRVQ 313
Db        195 ETVTHQFAV----KRYVLPKFEVTVNAPQTVTISDDEF-------QVDVCAKYNFGQPVQ 243

Qy        314 NVSVLSICDFREDWQTWDCPKTQEHIEES-----RTGWAGGFFVSTPVFSYDAISYYKIF 368
Db        244 GETQIRVC--REYFSSSNCEKNENEICEQFIAQLENGCVSQ-IVNTKVFQLYRSGLFMTF 300

Qy        369 SDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQ----------DSL------FYSS 412
Db        301 -------HVAVIVTESGTVMQIS----EKTSVF-ITQLLGTVNFENMDTFYRRGISYFGT 348

Qy        413 NEFEE---YPGRRNIYRIS-----IGSY-----------------PPSKKCVTCHLRKE 446
Db        349 LKFSDPNNVPMVNKLLQLELNDEFIGNYTTDENGEAQFSIDTSDIFDPEFNLKATYVRPE 408

Qy        447 RC--------QYYTASFSDYAKYYALVCYGPGIPISTLHDGRTDQEIKILEENKELE-NA 497
Db        409 SCYLPSWLTPQYLDAHFL-VSRFYS---------------RTNSFLKIVPEPKQLECNQ 451

Qy        498 LKNIQLPKEEIKKLEVDE--ITLWYKMILPPQFDRSKKYPLLIQVYGG----PCSQSVRS 551
Db        452 QKVVTVHYSLNSEAYEDDSNVKFFYLMMVKGAILLSGQKEIRNKAWNGNFSFPISISADL 511

Qy        552 VFAVNWISYLASKEGMVIALVDGRGTAFQGDK---------------------------- 583
Db        512 APAAVLFVYTLHPSGEIVA----DSVRFQVDKCFKHKVNIKFSNEQGLPGSNASLCLQAA 567

Qy        584 -----LLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASG- 637
Db        568 PVLFCALRAVDRNVLLLKSEQQLSAESVYNMVPSIE---------PYGYFYHGLNLDDGK 618

Qy        638 -------TGLFKCGIAVAPVSSW---EYYASVYTERFMGLPTKDDNLEHYKNSTVMARAE 687
Db        619 EDPCIPQRDMFYNGLYYTPVSNYGDGDIYNIV---RNMGLKV-FTNLHYRKPEKIMVQCV 674

Qy        688 YFRNVDYLLIHGTADDNVHFQNSAQIAKALV-NAQVDFQAMWYSD----QNHGLSGLS 740
Db        675 VFR----LELH--VASGIRGENADYVEQAIIQTVRTNFPETWMWDLVSVDSSGSANLS 726


RESULT 10
US-10-453-372-1092
; Sequence 1092, Application US/10453372
; Publication No. US20060003323A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Alsobrook, et al.
;  TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
;  FILE REFERENCE: 21402-589 A
;  CURRENT APPLICATION NUMBER: US/10/453,372
;  CURRENT FILING DATE:  2003-06-03
;  PRIOR APPLICATION NUMBER: 09/789390
;  PRIOR FILING DATE: 2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 09/823187
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/195792
;  PRIOR FILING DATE: 2000-03-10
;  PRIOR APPLICATION NUMBER: 09/839446
;  PRIOR FILING DATE: 2001-03-19
;  PRIOR APPLICATION NUMBER: 60/199476
;  PRIOR FILING DATE: 2000-03-25
;  PRIOR APPLICATION NUMBER: 09/863776
;  PRIOR FILING DATE: 2001-05-23
;  PRIOR APPLICATION NUMBER: 60/208263
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: 09/939398
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: 60/227800
;  PRIOR FILING DATE: 2000-08-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1609
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 1092
;   LENGTH: 1593
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-453-372-1092

  Query Match             2.6%;  Score 108;  DB 6;  Length 1593;
  Best Local Similarity   19.0%;  Pred. No. 1.3;
  Matches  148;  Conservative  109;  Mismatches  285;  Indels  236;  Gaps   41;

Qy          96 KSVNASNYGLSPDR------------QFVYLESDYSKLWRYSYTATYYIYDLSNGEFVRG 143
               ||:  ||  |                | | |  :| :|   :  :  |  :| :| : ||
Db         209 KSLPGSNIDLQVSAASNSLCALWAVDQSVLLLRNYGQLSAQTVYSQLYSREL-HGYYFRG 267

Qy         144 NELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRP-----GDPPFQITFNGRENKIFNGIPD 198
                 |   ::  |          ::  | ||    |      |   :  :  :  : |:||
Db         268 LNLEDGLKVPCLED-----EHILYNGIY--YTPAWADFGKDGYDLVKDPQNNRIFQ---- 316

Qy         199 WVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYGDEQYPRTINIPYPKAGAK 258
                                 ::          :|:   :||  |   :     |  ::
Db         317 ------------------RQNVTSFRNITQLSFQLISEPMFGD--YWIVV-----KRNSR 351

Qy         259 NPVVRIFIIDTTYPAYVGPQ-EVPV--PAMIASSDYYFSWLTWVTDERVCLQWL--KRVQ 313
                 |   | :      || |: || |  |  :  ||  |           || ::   : ||
Db         352 ETVTHQFAV----KRYVLPKFEVTVNAPQTVTISDDEF-------QVDVCAKYNFGQPVQ 400

Qy         314 NVSVLSICDFREDWQTWDCPKTQEHIEES-----RTGWAGGFFVSTPVFSYDAISYYKIF 368
                 : : :|  || : : :| | :  | |         |      |:| ||       :  |
Db         401 GETQIRVC--REYFSSSNCEKNENEICEQFIAQLENGCVSQ-IVNTKVFQLYRSGLFMTF 457

Qy         369 SDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQ----------DSL------FYSS 412
                    |:  |       :||:    |  ::| :||          |:       :: :
Db         458 -------HVAVIVTESGTVMQIS----EKTSVF-ITQLLGTVNFENMDTFYRRGISYFGT 505

Qy         413 NEFEE---YPGRRNIYRIS-----IGSY-----------------PPSKKCVTCHLRKE 446
                :| :    |    : ::      ||:|                 |       ::| |
Db         506 LKFSDPNNVPMVNKLLQLELNDEFIGNYTTDENGEAQFSIDTSDIFDPEFNLKATYVRPE 565

Qy         447 RC--------QYYTASFSDYAKYYALVCYGPGIPISTLHDGRTDQEIKILEENKELE-NA 497
                 |        ||  | |   :::|:                ||:  :||: | |:|| |
Db         566 SCYLPSWLTPQYLDAHFL-VSRFYS---------------RTNSFLKIVPEPKQLECNQ 608

Qy         498 LKNIQLPKEEIKKLEVDE--ITLWYKMILPPQFDRSKKYPLLIQVYGG----PCSQSVRS 551
                 | : :      :   |:  :  :| |::      | :  :  : : |    | | |
```

```
Db         609 QKVVTVHYSLNSEAYEDDSNVKFFYLMMVKGAILLSGQKEIRNKAWNGNFSFPISISADL 668

Qy         552 VFAVNWISYLASKEGMVIALVDGRGTAFQGDK---------------------------- 583
                 |      |      | ::|         || ||
Db         669 APAAVLFVYTLHPSGEIVA----DSVRFQVDKCFKHKVNIKFSNEQGLPGSNASLCLQAA 724

Qy         584 -----LLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASG- 637
                    | || | : : : | |::|    :  :  |:           || :   | | |
Db         725 PVLFCALRAVDRNVLLLKSEQQLSAESVYNMVPSIE---------PYGYFYHGLNLDDGK 775

Qy         638 -------TGLFKCGIAVAPVSSW---EYYASVYTERFMGLPTKDDNLEHYKNSTVMARAE 687
                        :|  |:   |||::   : |  |   | |||      || : |   :| :
Db         776 EDPCIPQRDMFYNGLYYTPVSNYGDGDIYNIV---RNMGLKV-FTNLHYRKPEKIMVQCV 831

Qy         688 YFRNVDYLLIHGTADDNVHFQNSAQIAKALV-NAQVDFQAMWYSD----QNHGLSGLS 740
                 ||    | :|      :  :|:  : :|::   : :|   |  |     : | : ||
Db         832 VFR----LELH--VASGIRGENADYVEQAIIQTVRTNFPETWMWDLVSVDSSGSANLS 883


RESULT 11
US-11-151-601-23
; Sequence 23, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  Meyers, Rachel E.
;  APPLICANT:  Olandt,  Peter J.
;  APPLICANT:  Kapeller-Libermann, Rosana
;  APPLICANT:  Curtis, Rory A. J.
;  APPLICANT:  Williamson, Mark
;  APPLICANT:  Weich, Nadine
;  TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
;  TITLE OF INVENTION:  AND PROTEASE FAMILY MEMBERS AND USES THEREOF
;  FILE REFERENCE: MPI00-054P1RCP1OMNIDV1M
;  CURRENT APPLICATION NUMBER: US/11/151,601
;  CURRENT FILING DATE:  2005-06-13
;  PRIOR APPLICATION NUMBER: US 10/170,789
;  PRIOR FILING DATE: 2002-06-13
;  PRIOR APPLICATION NUMBER: US 09/797,039
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: PCT/US01/06525
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: US 60/186,061
;  PRIOR FILING DATE: 2000-02-29
;  PRIOR APPLICATION NUMBER: US 09/882,166
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: PCT/US01/19269
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: US 60/212,078
;  PRIOR FILING DATE: 2000-06-15
;  PRIOR APPLICATION NUMBER: US 09/934,406
;  PRIOR FILING DATE: 2001-08-21
;  PRIOR APPLICATION NUMBER: PCT/US01/26052
;  PRIOR FILING DATE: 2001-08-21
;  PRIOR APPLICATION NUMBER: US 60/226,740
;  PRIOR FILING DATE: 2000-08-21
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 45
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23
;   LENGTH: 710
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: consensus sequence
US-11-151-601-23

  Query Match             2.6%;  Score 105.5;  DB 7;  Length 710;
  Best Local Similarity   18.9%;  Pred. No. 0.66;
  Matches  142;  Conservative  107;  Mismatches  260;  Indels  243;  Gaps   41;

Qy          35 EENTMRALTLKDILNGTFSYKTFFPNWISGQEYLHQSADNNIVLYNIETGQSYTILSNRT 94
                |:  :| |  |: :   : |  :  ::  |: | :          ||       | | |:
```

```
Db         55 EQCPIRGL-YKERMTELYDYPKYSCHFKKGKRYFY--------FYN-------TGLQNQR 98

Qy         95 MKSVNASNYG-----LSPDRQFVYLESDYSKLWR---YSYTATYYIYDLSNG-------E 139
              :  |  |  |     | |:     |  | :   |   :|    |: | ||         :
Db         99 VLYVQDSLEGEARVFLDPN----ILSDDGTVALRGYAFSEDGEYFAYGLSASGSDWVTIK 154

Qy        140 FVR---GNELP---RPIQYLC--WSPVGSKLAYVYQNNIYLKQRPGDPPFQITFNGRENK 191
              |::     |||     ::: |  |:  |  : |    | |  |: |       |      |
Db        155 FMKVDGAKELPDVLERVKFSCMAWTHDGKGMFY----NSY-PQQDGKSDGTETSTNLHQK 209

Qy        192 IFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYGDEQYPRT--IN 249
              ::  :      |::|  ::       |::  || :|    |:     | :  |    :
Db        210 LYYHVLGTDQSEDILCAEF----PDEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCD 265

Qy        250 IPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVTDERVCLQWL 309
              :    :|    :  : :||      |               ||       ||:|     :
Db        266 LQQESSGIAGILKWVKLIDNFEGEY--------------DY-------VTNEGTVFTFK 303

Qy        310 KRVQ--NVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGPFVSTPVFSYDAISYYKI 367
                |  |  |::| ||       ||          |||:  |      |  |     | : :  |
Db        304 TNRQSPNYRVINI-DF------WD-------PEESK--WK----VLVPEHEKDVLEW--I 341

Qy        368 FSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPGRRNIYRI 427
              :  :  : |: | |:| :|:            :|  :|                : :
Db        342 ACVRSNFLVLCYLHD-VKNILQLHD----------LTTGALL-------------KTFPL 377

Qy        428 SIGS---YPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCY-GPGIPISTLHDGRTDQ 483
              :||   |   ||                    |   :|    :  |||
Db        378 DVGSIVGYSGQKK------------------DTEIFYQFTSFLSPGI------------ 406

Qy        484 EIKILEENKELENALKNIQLPKEEIKKLEVDEITL---------WYKMILP-------PQ 527
                        : :  | |||::     |:|:           ::  |       |
Db        407 -------------IYHCDLTKEELEPRVFREVTVKGIDASDYQTVQIFYPSKDGTKIPM 452

Qy        528 FDRSKK-------YPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQ 580
              |  .||       :|  :  |||  : |:   ::|: : ::    | ::|: : ||
Db        453 FIVHKKGIKLDGSHPAFLYGYGG-FNISITPNYSVSRLIFVRHMGG-ILAVANIRGGGEY 510

Qy        581 GDKLLYAVYRKLGVYE-----VEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALA 635
              |:      : | |:        :|    |     |: |:   ||: | | | || : :
Db        511 GE-----TWHKGGILANKQNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVAACAN 565

Qy        636 SGTGLFKCGIAVAPV-----------SSW--EYYASVYTERFMGLPTKDDNLEHYK--- 678
                 || | ||   |              :|  :|  |                |: :|::
Db        566 QRPDLFGCVIAQVGVMDMLKFHKYTIGHAWTTDYGCS------------DSKQHFEWLV 612

Qy        679 -----NSTVMARAEYFRNVDYLLIHGTADDNV 705
                  ::  :  |:  :    ||:    || |
Db        613 KYSPLHNVKLPEADDIQYPSMLLLTADHDDRV 644
```

```
RESULT 12
US-11-110-851-65
; Sequence 65, Application US/11110851
; Publication No. US20050260219A1
; GENERAL INFORMATION:
;  APPLICANT: LAVAZEC, Catherine
;  APPLICANT:  BOURGOUIN, Catherine
;  APPLICANT:  BOUDIN, Christian
;  APPLICANT:  BONNET, Sarah
;  TITLE OF INVENTION: CARBOXYPEPTIDASES B FROM ANOPHELES GAMBIAE.  COMPOSITIONS
;  TITLE OF INVENTION:  COMPRISING THEM, VACCINE APPLICATIONS AND USE AS THERAPEUTICAL
;  TITLE OF INVENTION:  TARGETS
;  FILE REFERENCE: 268127US0
;  CURRENT APPLICATION NUMBER: US/11/110,851
;  CURRENT FILING DATE:  2005-04-21
;  PRIOR APPLICATION NUMBER: US 60/564,240
;  PRIOR FILING DATE: 2004-04-22
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 65
;   LENGTH: 433
;   TYPE: PRT
;   ORGANISM: Anopheles gambiae
US-11-110-851-65

  Query Match             2.6%;  Score 104.5;  DB 7;  Length 433;
  Best Local Similarity   21.4%;  Pred. No. 0.39;
  Matches   63;  Conservative   31;  Mismatches  104;  Indels   97;  Gaps   14;

Qy         81 IETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESD-YSKLWRYSYTATYYIYDLSNGE 139
              :: |:|:    |||||  |  | ||  | |  |:||   ::: |    | || :  |   |
Db        155 LDAGRSH---QNRTMKGVKLS-YG--PGRPGVFLEGGIHAREWISPATVTYILNQLLTSE 208

Qy        140 FVRGNELPRPIQYLCW---SPVGSKLAYVYQ-NNIYLKQRPGDPPF-------------- 181
                :   |      :  :   :| |   || :| | :: | |    ||
Db        209 DAKVRALAEKFDWYVFPNANPDG--YAYTFQVNRLWRKTRKAYGPFCYGADPNRNWDFHW 266

Qy        182 --QITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYY 239
                | | |   :  ::|     :  |    :  |       ||  ||  |:
Db        267 AEQGTSNNACSDTYHG--SEAFSEVETRSLAAFVEKLRGKLGAYIAFH------------ 312

Qy        240 GDEQYPRTINIPYPKAGAK-------NPVVRIF----------------IIDTTYPAYVG 276
                  | : :  ||   ||         | :                  :||  |||
Db        313 ---SYSQLLLFPYGDTGAHCGNHQDLNEIAEATVKSLAKRYGTQYKYGNVIDAIYPA--- 366

Qy        277 PQEVPVPAMIASSDYYFSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWD 331
                          ||                  : |    |:| :    : | |   |:
Db        367 ---------SGSS----------------VDWSYGAQDVKIAYTYELRPDGDAWN 396
```

```
RESULT 13
US-10-876-787-6
; Sequence 6, Application US/10876787
; Publication No. US20050287535A1
; GENERAL INFORMATION:
;  APPLICANT: McGrath, Kevin P.
;  APPLICANT:  Kimberly-Clark Worldwide, Inc.
;  TITLE OF INVENTION: BIOMARKERS FOR WOUND HEALING
;  FILE REFERENCE: 1443.156US1
;  CURRENT APPLICATION NUMBER: US/10/876,787
;  CURRENT FILING DATE:  2004-06-24
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 2671
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-876-787-6

  Query Match             2.5%;  Score 103;  DB 6;  Length 2671;
  Best Local Similarity   18.5%;  Pred. No. 7.4;
  Matches  156;  Conservative  103;  Mismatches  250;  Indels  336;  Gaps   45;

Qy         35 EENTMRALTLKDILNGTFSYKTFFPNWISGQEYLHQSADN-----NIVLYNIETGQ---- 85
              |:| || :|| |       |:    | |       |  : ||      ::|  :  ||
Db        144 EKNAMR-VTL-DATGNEGSWLFIQPFW-----KLRSNGDNVVVGDKVILNPVNAGQPLHA 196

Qy         86 -SYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSKLWRYSYTATYYIYDLSNGEFVRGN 144
               :| :  |   | ||: |   |                |: :    :  :     | ::|
Db        197 SNYELSDNAGCKEVNSVNCNTS--------------WKINLFMQFRDH---LEEVLKGG 238

Qy        145 ELPRPI-----QYLCWSPVGSKLAYVYQNNIYLKQRPGDPPFQITFNG--------RENK 191
              :: |        ::|        ||        ::|:           | :           :
Db        239 DVVRLFHAEQEKFLTCDEYKGKL------QVFLRTTLRQSATSATSSNALWEVEVVHHDP 292

Qy        192 IFNGIPDW--VYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYGDEQYPRTIN 249
                  |   |  :|  : |||           | :|| || |           || ||    |:
Db        293 CRGGAGHWNGLYRFKHLAT---------GNYLA-AEENP--------SYKGDASDPKAAG 334

Qy        250 IPYP------KAGAK-----------NPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYY 292
              :          || |           | :  :| :|                |  :  :| :
Db        335 MGAQGRTGRRNAGEKIKYCLVAVPHGNDIASLFELD-------------PTTLQKTDSF 380

Qy        293 FSWLTWVTDERVCLQ-WLKRVQNVSVLSICDFRED----WQTWDCPKTQEHIEESRTGWA 347
                  ::|    :|   |::   || :    |  |:          ||      :| :  :|
```

```
Db        381 VPRNSYVRLRHLCTNTWIQST-NVPI----DIEEERPIRLMLGTCP-----TKEDKEAFA 430

Qy        348 GGFFVSTPV-------FSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINI 400
                 || ||        |: ||| |              :    ||   ::   |
Db        431 ---IVSVPVSEIRDLDFANDASS--------------MLASAVE---KLNEG------- 462

Qy        401 FRVTQDSLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAK 460
                  | | |:      || : ::                                |
Db        463 --------FISQND------RRFVIQL---------------------------LED--- 478

Qy        461 YYALVCYGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEV------- 513
                 || :   :|    ::|:   :| : : |:| :  ::   : |:    |:|
Db        479 ---LVFFVSDVP----NNGQNVLDIMVTKPNRERQKLMREQNILKQVFGILKVPFREKGG 531

Qy        514 -------DEIT-----------LWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFA 554
                     :|::           | |:::   | |  |    : : :|   ||
Db        532 EGPLVRLEELSDQKNAPYQHMFRLCYRVLRYSQEDYRKNQEHIAKQFGMMQSQ------- 584

Qy        555 VNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEMGFID 614
                 | |    |  : ||:          |||     |    |||  :: |||  |  |:|
Db        585 ---IGYDILAEDTITALL------HNNRKLLEKHITKT---EVETFVSLVRKNREPRFLD 632

Qy        615 EKRIAIWGWSYGGYVSSLALASGTG-------LFKCGI-----------AVAPV----SS 652
                      |:| | :::          : || :            : ||     |
Db        633 ------------YLSDLCVSNHIAIPVTQELICKCVLDPKNSDILIRTELRPVKEMAQS 679

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVM----ARA---------EYFR--------- 690
                 || :  |:|  : |   | | ||:: |        |||           |:|
Db        680 HEYLSIEYSEEEVWLTWTDKNNEHHEKSVRQLAQEARAGNAHDENVLSYYRYQLKLFARM 739

Qy        691 ----------------NVDYLLIHGTADDNVHFQNSAQIAKALVNAQVD----------- 723
                             || |:    ||: : |   |     ::  ||
Db        740 CLDRQYLAIDEISQQLGVD-LIFLCMADEMLPFDLRASFCHLMLHVHVDRDPQELVTPVK 798

Qy        724 FQAMW 728
                |  :|
Db        799 FARLW 803
```

RESULT 14
US-10-467-657-1554
; Sequence 1554, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1554
; LENGTH: 621
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1554

Query Match 2.5%; Score 101.5; DB 6; Length 621;
Best Local Similarity 23.3%; Pred. No. 1.2;
Matches 55; Conservative 28; Mismatches 97; Indels 56; Gaps 11;

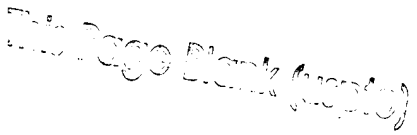
```
Qy        535 PLLIQVYGGPCSQSVRSVFAVNWI-SYLAS------KEGMVIALVDGRGTAFQGDKLLYA 587
                 | |:  |||          | :  :  || |      :||    | : ||    | :   |
Db        393 PTLVYAYGG---------FGIPELPHYLGSVGKYWLEEGNAFVLANIRGGGEFGPRWHQA 443

Qy        588 VYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALA---SGTGLFKCG 644
                     |:| :  ||   | |    | | : | | || ::: |        |   |
Db        444 AQGISKHKSVDDLLAVVRDLSERGMSSPKHIGLQGGSNGGLITAAAFVREPQSIGALVCE 503

Qy        645 IAVAPV---------SSWEYYASVYTERFMGLPTKDD-------NLEHYKNSTVMARAEY 688
                  : :  :          |||        |: : | | | :         |  | | :
Db        504 VPLTDMIRYPLLSAGSSW-------TDEY-GNPQKYEACKRRLGELSPYHNLS------- 548

Qy        689 FRNVDY---LLIHGTADDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLST 741
                     :||    |:    :|| ||  :: :    |   :    |:  ||    | :| |
Db        549 -DGIDYPPALITTSLSDDRVHPAHALKFYAKL--RETSPQSWLYSPDGGGHTGNGT 601
```

RESULT 15
US-10-878-556A-125
; Sequence 125, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 125
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: sw_hum/estd_human
; DATABASE ENTRY DATE: 1989-07-01
US-10-878-556A-125

Query Match 2.4%; Score 100; DB 6; Length 282;
Best Local Similarity 23.7%; Pred. No. 0.5;
Matches 55; Conservative 34; Mismatches 83; Indels 60; Gaps 11;

```
Qy        509 KKLEVDEITLWYKM----ILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWIS----- 559
                 |  | | : |  ||      |||: : : | | |  : |  |::        |:||
Db         17 KVFEHDSVELNCKMKFAVYLPPKAE-TGKCPALYWLSGLTCTEQ-------NFISKSGYH 68

Qy        560 YLASKEGMVIALVD----------------GRGTAFQGD------KLLYAVYRKLGVYE 596
                 ||: |:|:   |                | |  | |      |  | :|     |
Db         69 QSASEHGLVVIAPDTSPRGCNIKGEDESWDFGTGAGFYVDATEDPWKTNYRMYS----YV 124

Qy        597 VEDQITAVRKFIEMGF-IDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS--- 652
                  |:    : : |   | :| :|::|:| | ||: : :      | :|   | ||: :
Db        125 TEE----LPQLINANFPVDPQRMSIFGHSMGGHGALICALKNPGKYKSVSAFAPICNPVL 180

Qy        653 --WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTAD 702
                    |   |      | |    | :     ::| : ::     :| |:  |  |
Db        181 CPWGKKA------FSGYLGTDQSKWKAYDATHLVKSYPGSQLDILIDQGKDD 226
```

Search completed: January 28, 2006, 00:47:02
Job time : 37 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2006, 00:26:19 ; Search time 40 Seconds
(without alignments)
1828.118 Million cell updates/sec

Title: US-10-681-659-1
Perfect score: 4086
Sequence: 1 MKTWVKIVFGVATSAVLALL.....TNHLYTHMTHFLKQCFSLSD 760

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3822.5	93.6	759	2	I38593	fibroblast activat
2	2192	53.6	792	1	A39914	dipeptidyl-peptida
3	2186.5	53.5	766	1	CDHU26	dipeptidyl-peptida
4	2179.5	53.3	760	1	S23752	dipeptidyl-peptida
5	1177	28.8	803	2	A41793	dipeptidyl aminope
6	1161	28.4	803	2	I68600	dipeptidyl aminope
7	1161	28.4	865	2	I54331	dipeptidyl aminope
8	864.5	21.2	793	2	T41703	dipeptidyl aminope
9	841	20.6	711	2	S66261	X-Pro dipeptidyl-p
10	814.5	19.9	818	1	A30107	dipeptidyl aminope
11	774.5	19.0	829	2	T19514	hypothetical prote
12	744	18.2	779	2	T25173	hypothetical prote
13	739	18.1	799	2	T25174	hypothetical prote
14	721	17.6	931	2	A49737	dipeptidyl aminope
15	622.5	15.2	738	2	A87516	dipeptidyl peptida
16	609.5	14.9	741	2	JC5142	X-Pro dipeptidyl-p
17	596	14.6	743	2	T37700	probable dipeptidy
18	482	11.8	931	2	T32919	hypothetical prote
19	314	7.7	795	2	F82858	dipeptidyl-peptida
20	257.5	6.3	657	2	E70025	probable acylamino
21	232	5.7	642	2	C71137	hypothetical prote
22	225	5.5	631	2	H75007	probable acylamino
23	218	5.3	764	2	JC8016	acylaminoacyl-pept
24	210.5	5.2	632	2	E75057	peptidase PAB1418
25	209.5	5.1	622	2	F71174	hypothetical prote
26	194	4.7	642	2	D87379	prolyl oligopeptid
27	189.5	4.6	591	2	H72474	probable acylamino
28	178.5	4.4	674	2	B84381	acylaminoacyl-pept
29	177	4.3	683	2	E87495	prolyl oligopeptid
30	172.5	4.2	721	2	T09631	probable acylamino
31	170	4.2	659	2	F72568	probable acylamino
32	169.5	4.1	761	2	S44807	F44B9.1 protein -
33	166	4.1	582	2	D72636	probable acylamino
34	164.5	4.0	732	1	S07624	acylaminoacyl-pept
35	159.5	3.9	709	2	B82580	alanyl dipeptidyl
36	158.5	3.9	732	1	JC4655	acylaminoacyl-pept
37	158	3.9	745	2	T33751	hypothetical prote
38	157.5	3.9	655	2	E75551	probable acyl-pept
39	155.5	3.8	676	2	C97775	acylamino-acid-rel
40	154.5	3.8	819	2	B87580	conserved hypothet
41	152	3.7	955	2	F84914	hypothetical prote
42	150	3.7	637	2	S75772	hypothetical prote
43	149.5	3.7	606	2	T35378	probable peptide h
44	149.5	3.7	618	2	D87651	prolyl oligopeptid
45	149	3.6	572	2	F72455	probable acylamino

ALIGNMENTS

RESULT 1
I38593
fibroblast activation protein-alpha - human
N;Alternate names: FAP-alpha
C;Species: Homo sapiens (man)
C;Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 29-Aug-1997
C;Accession: I38593
R;Scanlan, M.J.; Raj, B.; Calvo, B.; Garin-Chesa, P.; Sanz-Moncasi, M.P.; Healey, J.; O
Proc. Natl. Acad. Sci. U.S.A. 91, 5657-5661, 1994
A;Title: Molecular cloning of fibroblast activation protein alpha, a member of the seri
A;Reference number: I38593; MUID:94261645; PMID:7911242
A;Accession: I38593
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-759 <RES>
A;Cross-references: UNIPARC:UPI0000175B7D; EMBL:U09278; NID:g507749; PID:g507750
C;Genetics:
A;Gene: GDB:FAP
A;Cross-references: GDB:374184; OMIM:600403
A;Map position: 2q23-2q23
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: transmembrane protein

Query Match 93.6%; Score 3822.5; DB 2; Length 759;
Best Local Similarity 92.9%; Pred. No. 1.2e-257;
Matches 719; Conservative 7; Mismatches 19; Indels 29; Gaps 4;

```
Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              |||||||||||||||||||||||| |||||||||||||||||| ||||||||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
```

```
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEY---YA 657

Db        601 ITAVRKFIEMGFIDEKRIAIWGWSY-----EIRFITGPCIW----------NWSFQMWYS 645

Qy        658 S----------VYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVH 706

Db        646 SGSSLQLGILRVCLHREIHGSPNKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVH 705

Qy        707 FQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760

Db        706 FQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 759
```

RESULT 2
A39914
dipeptidyl-peptidase IV (EC 3.4.14.5), membrane-bound form precursor - rat
N;Alternate names: GP110; membrane glycoprotein 110K; OX-61
N;Contains: dipeptidyl-peptidase IV, soluble form
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A39914; A33315; B33315; A60730; A42203; S38949; A31781
R;Hong, W.; Doyle, D.
Proc. Natl. Acad. Sci. U.S.A. 84, 7962-7966, 1987
A;Title: cDNA cloning for a bile canaliculus domain-specific membrane glycoprotein of ra
A;Reference number: A39914; MUID:88068516; PMID:3479775
A;Accession: A39914
A;Molecule type: mRNA
A;Residues: 1-792 <HON>
A;Cross-references: UNIPROT:P14740; UNIPARC:UPI000017098A; GB:J02997; NID:g204463; PIDN:
R;Ogata, S.; Misumi, Y.; Ikehara, Y.
J. Biol. Chem. 264, 3596-3601, 1989
A;Title: Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA ar
A;Reference number: A33315; MUID:89123496; PMID:2563382
A;Accession: A33315
A;Molecule type: mRNA
A;Residues: 1-37,'A',39-182,'I',184-331,'T',333-351,'C',353-393,'V',395-561,'L',563-623,
A;Cross-references: UNIPARC:UPI0000129841; GB:J04591; NID:g203973; PIDN:AAA41096.1; PID:
A;Note: the authors translated the codon GCG for residue 38 as Arg, ACC for residue 332
A;Accession: B33315
A;Molecule type: protein
A;Residues: 1-20;35-54;427-443;505-509;511-520;530-538;593-600;602-608;618-627 <HO2>
A;Cross-references: UNIPARC:UPI0000172A2E; UNIPARC:UPI0000172A30; UNIPARC:UPI0000172A31;
A36; UNIPARC:UPI0000172A37
R;McCaughan, G.W.; Wickson, J.E.; Creswick, P.F.; Gorrell, M.D.
Hepatology 11, 534-544, 1990
A;Title: Identification of the bile canalicular cell surface molecule GP110 as the ectop
quence.
A;Reference number: A60730; MUID:90228896; PMID:1970322
A;Accession: A60730
A;Molecule type: protein
A;Residues: 28-47,'XX',50-53,55-58 <MCC>
A;Cross-references: UNIPARC:UPI0000172A38
R;Ogata, S.; Misumi, Y.; Tsuji, E.; Takami, N.; Oda, K.; Ikehara, Y.
Biochemistry 31, 2582-2587, 1992
A;Title: Identification of the active site residues in dipeptidyl peptidase IV by affini
A;Reference number: A42203; MUID:92190188; PMID:1347701
A;Accession: A42203
A;Molecule type: protein
A;Residues: 'R',625-630,'X',632-648 <OG2>
A;Cross-references: UNIPARC:UPI0000172A39
R;Iwaki-Egawa, S.; Watanabe, Y.; Fujimoto, Y.
Biol. Chem. Hoppe-Seyler 374, 973-975, 1993
A;Title: N-terminal amino acid sequence of the 60-kDa protein of rat kidney dipeptidyl
A;Reference number: S38949; MUID:94128239; PMID:7905271
A;Accession: S38949
A;Status: preliminary
A;Molecule type: protein
A;Residues: 281-302 <IWA>
A;Cross-references: UNIPARC:UPI0000172A3A
R;Hong, W.; Doyle, D.
J. Biol. Chem. 263, 16892-16898, 1988
A;Title: Membrane orientation of rat gp110 as studied by in vitro translation.
A;Reference number: A31781; MUID:89034185; PMID:3182821
A;Accession: A31781
A;Molecule type: mRNA
A;Residues: 1-40 <HO3>
A;Cross-references: UNIPARC:UPI0000172A3B
C;Comment: This protein is localized to the bile canaliculus, which is the apical domai
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidylpeptide hydrolase; glycoprotein; homodimer; liver; serine protein
F;1-792/Product: dipeptidyl-peptidase, membrane-bound form #status experimental <MATM>
F;1-28/Domain: signal sequence #link MATS #status experimental <SIG>
F;1-6/Domain: intracellular #status predicted <INT>
F;7-28/Domain: transmembrane #status predicted <TMN>
F;29-792/Domain: extracellular #status predicted <EXT>
F;29-34/Domain: propeptide #link MATS #status experimental <PRO>
F;35-792/Product: dipeptidyl-peptidase, soluble form #status experimental <MATS>
F;83,90,148,217,227,319,521,686/Binding site: carbohydrate (Asn) (covalent) #status pre
F;631/Active site: Ser #status experimental
F;709,741/Active site: Asp, His #status predicted

Query Match 53.6%; Score 2192; DB 1; Length 792;
Best Local Similarity 52.1%; Pred. No. 3e-144;
Matches 403; Conservative 150; Mismatches 199; Indels 22; Gaps 11;

```
Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEE---NTMRALTLKDILNGTFSYKT 56

Db          1 MKTPWKVLLGLLGVAALVTIITVPVVLL-----NKDEAAADSRRTYTLADYLKNTFRVKS 55

Qy         57 FFPNWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMK--SVNASNYGLSPDRQFVYL 114

Db         56 YSLRWVSDSEYLYKQ-ENNILLFNAEHGNSSIFLENSTFEIFGDSISDYSVSPDRLFVLL 114

Qy        115 ESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQ 174

Db        115 EYNYVKQWRHSYTASYSIYDLNKRQLITEEKIPNNTQWITWSQEGHKLAYVWKNDIYVKI 174

Qy        175 RPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVI 234

Db        175 EPHLPSHRTTSTGKENVIFNGINDWVYEEEIFGAYSALWWSPNGTFLAYAQFNDTGVPLI 234

Qy        235 AYSYYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDT---TYPAYVGPQEVPVPAMIASS 289

Db        235 EYSFYSDESLQYPKTVWIPYPKAGAVNPTVKFFIVNTDSLSSTTTTIPMQITAPASVTTG 294

Qy        290 DYYFSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGG 349

Db        295 DHYLCDVAWVSEDRISLQWLRRIQNYSVMAICDYDKTNLVWNCPTTQEHIETSATGWVGR 354

Qy        350 FFVSTPVFSYDAISYYKIFSDKDGYKHI-HYIKD--TVENAIQITSGKWEAINIFRVTQD 406

Db        355 FRPAEPHFTSDGSSFYKIVSDKDGYKHICQFQKDRKPEQDCTFITKGAWEVISIEALTSD 414

Qy        407 SLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVC 466

Db        415 YLYYISNEYKEMPGGRNLYKIQLTDH-TNKKCLSCDLNPERCQYYSVSLSKEAKYYQLGC 473

Qy        467 YGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPP 526

Db        474 RGPGLPLYTLHRSTDQKELRVLEDNSALDKMLQDVQMPSKKLDFIVLNETRFWYQMILPP 533
```

```
Qy        527 QFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLY 586
              ||:||||||||  || |||||   : |  || :|||| | :::|  ||||: :||||::: 
Db        534 HFDKSKKYPLLIDVYAGPCSQKADAAFRFNWATYLASTENIIVASFDGRGSGYQGDKIMH 593

Qy        587 AVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIA 646
              |: ::||  ||||| | |:|::|||:| |::|||||||||||:|: | ||:|:|||||| 
Db        594 AINKRLGTLEVEDQIEAARQFLKMGFVDSKQVAIWGWSYGGYVTSMVLGSGSGVFKCGIA 653

Qy        647 VAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVH 706
              ||||| |||| ||||||:||||| :|||:||:|||||:||| |: |:||||||||||||| 
Db        654 VAPVSRWEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVH 713

Qy        707 FQNSAQIAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSLS 759
              || ||||:||||:| ||||||||:|::||: |  :  |:|:||:|||:|||||: 
Db        714 FQQSAQISKALVDAGVDFQAMWYTDEDHGIASSTAHQHIYSHMSHFLQQCFSLT 767
```

RESULT 3
CDHU26
dipeptidyl-peptidase IV (EC 3.4.14.5) - human
N;Alternate names: cell surface glycoprotein CD26; thymocyte-activating molecule (THAM)
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 23-Aug-1996 #text_change 09-Jul-2004
C;Accession: S24313; B42408; A42408; B61136; S59510; I56154; S59857; S15520
R;Misumi, Y.; Hayashi, Y.; Arakawa, F.; Ikehara, Y.
Biochim. Biophys. Acta 1131, 333-336, 1992
A;Title: Molecular cloning and sequence analysis of human dipeptidyl peptidase IV, a ser
A;Reference number: S24313; MUID:92329551; PMID:1352704
A;Accession: S24313
A;Molecule type: mRNA
A;Residues: 1-6,'I',8-766 <MIS>
A;Cross-references: UNIPROT:P27487; UNIPARC:UPI000016AE80; EMBL:X60708; NID:g35335; PIDN
R;Darmoul, D.; Lacasa, M.; Baricault, L.; Marguet, D.; Sapin, C.; Trotot, P.; Barbat, A.
J. Biol. Chem. 267, 4824-4833, 1992
A;Title: Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like colon cancer
 IV mRNA levels during cell differentiation.
A;Reference number: A42408; MUID:92165847; PMID:1347043
A;Accession: B42408
A;Molecule type: mRNA
A;Residues: 1-5,'R',7-436,'S',438-556,'I',558-662,'E',664-766 <DAR1>
A;Cross-references: UNIPARC:UPI0000052ACB; GB:M80536; NID:g181569; PIDN:AAA52308.1; PID:
A;Experimental source: intestine
A;Note: this sequence corresponds with the author's translation
A;Accession: A42408
A;Molecule type: mRNA
A;Residues: 1-5,'R',7-436,'S',438-556,'I',558-662,'E',664-711,'G',713-766 <DAR2>
A;Cross-references: UNIPARC:UPI0000172A2B; GB:M80536; NID:g181569
A;Note: sequence extracted from NCBI backbone (NCBIN:83986, NCBIP:83988); this sequence
R;Gorvel, J.P.; Ferrero, A.; Chambraud, L.; Rigal, A.; Bonicel, J.; Maroux, S.
Gastroenterology 101, 618-625, 1991
A;Title: Expression of sucrase-isomaltase and dipeptidylpeptidase IV in human small inte
A;Reference number: A61136; MUID:91317403; PMID:1677636
A;Accession: B61136
A;Molecule type: protein
A;Residues: 1-15,'X',17-22 <GOR>
A;Cross-references: UNIPARC:UPI0000172A2C
R;Boehm, S.K.; Gum Jr., J.R.; Erickson, R.H.; Hicks, J.W.; Kim, Y.S.
Biochem. J. 311, 835-843, 1995
A;Title: Human dipeptidyl peptidase IV gene promoter: tissue-specific regulation from a
A;Reference number: S59510; MUID:96067599; PMID:7487939
A;Accession: S59510
A;Molecule type: DNA
A;Residues: 1-31 <BOE>
A;Cross-references: UNIPARC:UPI000016B4A6; GB:S79876; NID:g1195574; PIDN:AAB35614.1; PID
R;Tanaka, T.; Camerini, D.; Seed, B.; Torimoto, Y.; Dang, N.H.; Kameoka, J.; Dahlberg, H
J. Immunol. 149, 481-486, 1992
A;Title: Cloning and functional expression of the T cell activation antigen CD26.
A;Reference number: I56154; MUID:92325476; PMID:1352530
A;Accession: I56154
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-436,'S',438-766 <TAN>
A;Cross-references: UNIPARC:UPI000004F7BF; GB:M74777; NID:g180082; PIDN:AAA51943.1; PID:
R;Abbott, C.A.; Baker, E.; Sutherland, G.R.; McCaughan, G.W.
Immunogenetics 40, 331-338, 1994
A;Title: Genomic organization, exact localization, and tissue expression of the human C
A;Reference number: S59857; MUID:95012454; PMID:7927537
A;Accession: S59857
A;Molecule type: DNA
A;Residues: 1-436,'S',438-766 <ABB>
A;Cross-references: UNIPARC:UPI000004F7BF; EMBL:U13734
C;Genetics:
A;Gene: GDB:DPP4
A;Cross-references: GDB:125239; OMIM:102720
A;Map position: 2q24.3-2q24.3
A;Introns: 2/3; 32/1; 65/1; 95/3; 122/3; 140/2; 164/3; 205/1; 258/3; 296/2; 341/3; 356/
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidylpeptide hydrolase; glycoprotein; homodimer; proteinase; transmemb
F;1-6/Domain: intracellular #status predicted <INT>
F;7-28/Domain: transmembrane #status predicted <TMN>
F;29-766/Domain: extracellular #status predicted <EXT>
F;85,92,150,219,229,281,321,520,685/Binding site: carbohydrate (Asn) (covalent) #status
F;630,708,740/Active site: Ser, Asp, His #status predicted

  Query Match   53.5%; Score 2186.5; DB 1; Length 766;
  Best Local Similarity 51.5%; Pred. No. 6.9e-144;
  Matches 395; Conservative 149; Mismatches 212; Indels   11; Gaps   7;

```
Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  |:: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db          1 MKTPWKVLLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
               |||  |||::  :|||:::| | | |   | | |       |  :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
              | ::||: |:|: |:||| |||||||: :   |||||||| |||||:||||::|:| ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :||||||| || |: |:::|   : |      ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
                :|| | ||: |||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||::  || ||:|:| :  |      |::| |  ||||||: |||  |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLIDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||    |: :::|||:|  |:  |:|:|:| :::  : ::|   ||:||||| ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||| | |:| :|||:| |||||||||||||||:|: | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| |||||:|||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718
```

```
Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |:||||:  ||||||||:|::||: |  :  |:||||:||:||||||
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765


RESULT 4
S23752
dipeptidyl-peptidase IV (EC 3.4.14.5) alpha chain - mouse
N;Alternate names: CD26 alpha subunit; THAM alpha subunit
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S23752; A46465; A56030
R;Marguet, D.; Bernard, A.M.; Vivier, I.; Darmoul, D.; Naquet, P.; Pierres, M.
J. Biol. Chem. 267, 2200-2208, 1992
A;Title: cDNA cloning for mouse thymocyte-activating molecule. A multifunctional ecto-di
A;Reference number: S23752; MUID:92129288; PMID:1370813
A;Accession: S23752
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-760 <MAR>
A;Cross-references: UNIPARC:UPI0000172A2D; EMBL:X58384
R;Vivier, I.; Marguet, D.; Naquet, P.; Bonicel, J.; Black, D.; Li, C.X.; Bernard, A.M.;
J. Immunol. 147, 447-454, 1991
A;Title: Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl peptidase
A;Reference number: A46465; MUID:91302787; PMID:1712807
A;Accession: A46465
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <VIV>
A;Cross-references: UNIPARC:UPI0000172A2E
A;Experimental source: M14.T thymoma cells, Swiss nu/nu
A;Note: sequence extracted from NCBI backbone (NCBIP:42236)
R;Bernard, A.M.; Mattei, M.G.; Pierres, M.; Marguet, D.
Biochemistry 33, 15204-15214, 1994
A;Title: Structure of the mouse dipeptidyl peptidase IV (CD26) gene.
A;Reference number: A56030; MUID:95092780; PMID:7999781
A;Accession: A56030
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 746-760 <BER>
A;Cross-references: UNIPARC:UPI0000172A2F; GB:U12620
C;Genetics:
A;Gene: CD26
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein
F;213,223,315,514,679/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;624,702,734/Active site: Ser, Asp, His #status predicted

  Query Match             53.3%;  Score 2179.5;  DB 1;  Length 760;
  Best Local Similarity   51.9%;  Pred. No. 2.1e-143;
  Matches  397;  Conservative  149;  Mismatches  206;  Indels   13;  Gaps    9;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              |||  |:: |:   | |  ::   ::  |:   :  :: |  :| | |  ||  |::
Db          1 MKTPWKVLLGLLGVAALVTIITVPIVLLSK-DEAAADSRRTYSLADYLKSTFRVKSYSLW 59

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              |:|  |||::  :|||:| | | | |    | | | :|     : :|||| || || :| |
Db         60 WVSDFEYLYKQ-ENNILLLNAEHGNSSIFLENSTFESF--GYHSVSPDRLFVLLEYNYVK 116

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
               ||:||||:| |||::  : :   ::|   |:: ||| | |||||::|:|||:|  |  |
Db        117 QWRHSYTASYNIYDVNKRQLITEEKIPNNTQWITWSPEGHKLAYVWKNDIYVKVEPHLPS 176

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              :||  | || |:||| |||||||:     |||||||  |||||:|||| :|:| ||:|
Db        177 HRITSTGEENVIYNGITDWVYEEEVFGAYSALWWSPNNTFLAYAQFNDTGVPLIEYSFYS 236

Qy        241 DE--QYPRTINIPYPKAGAKNPVVRIFIID---TTYPAYVGPQEVPVPAMIASSDYYFSW 295
              ||  |||:|: ||||||||  || |: ||::    :  :   | ::| || :|  |:|
Db        237 DESLQYPKTVWIPYPKAGAVNPTVKFFIVNIDSLSSSSSAAPIQIPAPASVARGDHYLCD 296

Qy        296 LTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTP 355
              : | |:||: ||||:|:|| ||::|||: :   ||:||  |:|:| | ||| | |  : |
Db        297 VVWATEERISLQWLRRIQNYSVMAICDYDKINLTWNCPSEQQHVEMSTTGWVGRFRPAEP 356

Qy        356 VFSYDAISYYKIFSDKDGYKHI-HYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNE 414
              |: |  |:||| ||||||||| |: ||  ::   || | || |:|  :| | |:| ||:
Db        357 HFTSDGSSFYKIISDKDGYKHICHFPKDK-KDCTFITKGAWEVISIEALTSDYLYYISNQ 415

Qy        415 FEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPIS 474
              ::| || ||:|:| :  :  : ||::| |  |||||  |||  |||| | |:|||:|:
Db        416 YKEMPGGRNLYKIQLTDH-TNVKCLSCDLNPERCQYYAVSFSKEAKYYQLGCWGPGLPLY 474

Qy        475 TLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKY 534
              |||     :|:::||:|  |:  |:::|:| :::  : ::|   ||:||||| ||:||||
Db        475 TLHRSTDHKELRVLEDNSALDRMLQDVQMPSKKLDFIVLNETRFWYQMILPPHFDKSKKY 534

Qy        535 PLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGV 594
              ||| || |||||   : | :|| :|||| | :::|  ||||: :||||:::|: |:||
Db        535 PLLLDVYAGPCSQKADASFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRLGT 594

Qy        595 YEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWE 654
              |||||| | |:|::|||:| ||:|||||||||||:|: | ||:|:|||||||||||| ||
Db        595 LEVEDQIEAARQFVKMGFVDSKRVAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSRWE 654

Qy        655 YYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLLIHGTADDNVHFQNSAQIA 714
              || ||||||:|||| :|||:||:|||||:|||:|: |:|||||||||||||| ||||:
Db        655 YYDSVYTERYMGLPIPEDNLDHYRNSTVMSRAEHFKQVEYLLLIHGTADDNVHFQQSAQIS 714

Qy        715 KALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              | ||:| ||||||||:|::||: |  :  |:|:||:|||:|||||
Db        715 KVLVDAGVDFQAMWYTDEDHGIASSTAHQHIYSHMSHFLQQCFSL 759


RESULT 5
A41793
dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short form, DPPX-S - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A41793
R;Wada, K.; Yokotani, N.; Hunter, C.; Doi, K.; Wenthold, R.J.; Shimasaki, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 197-201, 1992
A;Title: Differential expression of two distinct forms of mRNA encoding members of a dip
A;Reference number: A41793; MUID:92108018; PMID:1729689
A;Accession: A41793
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-803 <WAD>
A;Cross-references: UNIPROT:P42659; UNIPARC:UPI000002A83C; GB:M76429; NID:g408719; PIDN.
A;Note: sequence extracted from NCBI backbone (NCBIP:75138)
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidylpeptide hydrolase; glycoprotein
F;257,342/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             28.8%;  Score 1177;  DB 2;  Length 803;
  Best Local Similarity   32.6%;  Pred. No. 1.1e-73;
  Matches  257;  Conservative  174;  Mismatches  289;  Indels   68;  Gaps   23;

Qy         10 GVATSAVLALLVMC-------IVLRPSRVHNSEENTM---RALTLKDILNGTFSYKTFFP 59
              |:|  |:| :||:|        |:| |     :|:|::   : :|::|: :  |
Db         33 GIAI-ALLVILVICSLIVTSVILLTP-----AEDNSLSQKKKVTVEDLFSEDFKIHDPEA 86

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYS 119
              ||| :|:::      :::| |:|| |  ::  : ::|: |  | :||::      :
Db         87 KWISDKEFIYREQKGSVILRNVETNTSTVLIEGKKIESLRAIRYEISPDREYALFSYNVE 146

Qy        120 KLWRYSYTATYYIYDLSNGEFVRGNELPRP------IQYLCWSPVGSKLAYVYQNNIYLK 173
              ::::|||  | :  : :|:       |  |      :||  | | | :| ::::||||
Db        147 PIYQHSYTGYYVLSKIPHGD---PQSLDPPEVSNAKLQYAGWGPKGQQLIFIFENNIYYC 203

Qy        174 QRPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPV 233
                |    ::   |:|  |:||: ||:||||:| |  | ||||:|  ||||  ||: :||
Db        204 AHVGKQAIRVVSTGKEGVIYNGLSDWLYEEEILKTHIAHWWSPDGTRLAYATINDSRVPV 263

Qy        234 IAYSYYGDEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYF 293
```

```
Db        264 MELPTYTGSVYPTAKPYHYPKAGCENPSISLHVIGLNGPTH--DLEMTPPDDPRMREYYI 321

Qy        294 SWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVS 353

Db        322 TMVKWATSTKVAVNWLSRAQNVSILTLCDA----TTGVC--TKKHEDESE-AWLHR-QNE 373

Qy        354 TPVFSYDAISYYKIFSDKDG----YKHIHYIKDTV---------ENAIQITSGKWEAINI 400

Db        374 EPVFSKDGRKFFFVRAIPQGGQGKFYHI-----TVSSSQPNSSNDNIQSITSGDWDVTKI 428

Qy        401 --FRVTQDSLFYSSNEFEEYPGRRNIYRIS-IGSYPPSKKCVTCHLRKERCQYYTASFSD 457

Db        429 LSYDEKRSQIYFLST--EDLPRRRQLYSASTVGSF--NRQCLSCDL-VDNCTYFSASFSP 483

Qy        458 YAKYYALVCYGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEIT 517

Db        484 GADFFLLKCEGPGVPTVSVHNTTDKKKMFDLETNEHVQKAISDRQMPKVEYRKIETDDYN 543

Qy        518 LWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGT 577

Db        544 LPIQILKPATFTDTAHYPLLLVVDGTPGSQSVAEKFAVTWETVMVSSHGAVVVKCDGRGS 603

Qy        578 AFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLAL-AS 636

Db        604 GFQGTRLLHEVRRRLGSLEEKDQMEAVRVMLKEPYIDKTRVAVFGKDYGGYLSTYLLPAK 663

Qy        637 GTG---LFKCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVD 693

Db        664 GDGQAPVFSCGSALSPITDFKLYASAFSERYLGLHGLDNRA--YEMAKVAHRVSALEGQQ 721

Qy        694 YLLIHGTADDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLS-TNHLYTHMTHFL 752

Db        722 FLVIHATADEKIHFQHTAELITQLIKGKANYSLQIYPDESHYFSSAALQQHLHRSILGFF 781

Qy        753 KQCFSLSD 760

Db        782 VECFRIQD 789
```

RESULT 6
I68600
dipeptidyl aminopeptidase like protein - human
C;Species: Homo sapiens (man)
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C;Accession: I68600
R;Yokotani, N.; Doi, K.; Wenthold, R.J.; Wada, K.
Hum. Mol. Genet. 2, 1037-1039, 1993
A;Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-relat
A;Reference number: I54331; MUID:93372805; PMID:8103397
A;Accession: I68600
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-803 <RES>
A;Cross-references: UNIPARC:UPI000016A7F9; GB:M96860; NID:g306707; PIDN:AAA35761.1; PID:
C;Superfamily: dipeptidyl-peptidase IV

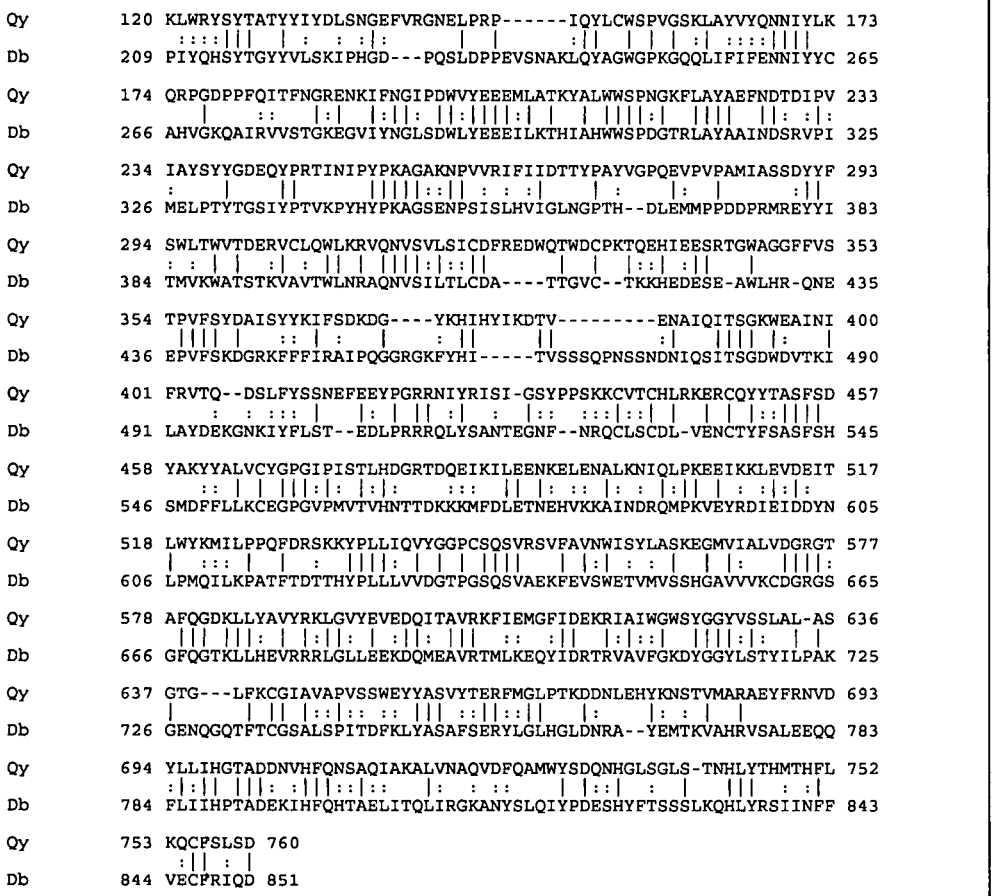
  Query Match   28.4%; Score 1161; DB 2; Length 803;
  Best Local Similarity   32.4%; Pred. No. 1.5e-72;
  Matches 255; Conservative 173; Mismatches 292; Indels   68; Gaps   23;

```
Qy         10 GVATSAVLALLVMC-------IVLRPSRVHNSEENTM---RALTLKDILNGTFSYKTFFP 59

Db         33 GIAI-ALLVILVICSLIVTSVILLTP-----AEDNSLSQKKKVTVEDLFSEDFKIHDPEA 86

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYS 119

Db         87 KWISDTEFIYREQKGTVRLWNVETNTSTVLIEGKKIESLRAIRYEISPDREYALFSYNVE 146

Qy        120 KLWRYSYTATYYIYDLSNGEFVRGNELPRP------IQYLCWSPVGSKLAYVYQNNIYLK 173

Db        147 PIYQHSYTGYYVLSKIPHGD---PQSLDPPEVSNAKLQYAGWGPKGQQLIFIFENNIYYC 203

Qy        174 QRPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPV 233

Db        204 AHVGKQAIRVVSTGKEGVIYNGLSDWLYEEEILKTHIAHWWSPDGTRLAYAAINDSRVPI 263

Qy        234 IAYSYYGDEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYF 293

Db        264 MELPTYTGSIYPTVKPYHYPKAGSENPSISLHVIGLNGPTH--DLEMMPPDDPRMREYYI 321

Qy        294 SWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVS 353

Db        322 TMVKWATSTKVAVTWLNRAQNVSILTLCDA----TTGVC--TKKHEDESE-AWLHR-QNE 373

Qy        354 TPVFSYDAISYYKIFSDKDG----YKHIHYIKDTV---------ENAIQITSGKWEAINI 400

Db        374 EPVFSKDGRKFFFIRAIPQGGRGKFYHI-----TVSSSQPNSSNDNIQSITSGDWDVTKI 428

Qy        401 FRVTQ--DSLFYSSNEFEEYPGRRNIYRISI-GSYPPSKKCVTCHLRKERCQYYTASFSD 457

Db        429 LAYDEKGNKIYFLST--EDLPRRRQLYSANTEGNF--NRQCLSCDL-VENCTYFSASFSH 483

Qy        458 YAKYYALVCYGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEIT 517

Db        484 SMDFFLLKCEGPGVPMVTVHNTTDKKKMFDLETNEHVKKAINDRQMPKVEYRDIEIDDYN 543

Qy        518 LWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGT 577

Db        544 LPMQILKPATFTDTTHYPLLLVVDGTPGSQSVAEKFEVSWETVMVSSHGAVVVKCDGRGS 603

Qy        578 AFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLAL-AS 636

Db        604 GFQGTKLLHEVRRRLGLLEEKDQMEAVRTMLKEQYIDRTRVAVFGKDYGGYLSTYILPAK 663

Qy        637 GTG---LFKCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVD 693

Db        664 GENQGQTFTCGSALSPITDFKLYASAFSERYLGLHGLDNRA--YEMTKVAHRVSALEEQQ 721

Qy        694 YLLIHGTADDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLS-TNHLYTHMTHFL 752

Db        722 FLIIHPTADEKIHFQHTAELITQLIRGKANYSLQIYPDESHYFTSSSLKQHLYRSIINFF 781

Qy        753 KQCFSLSD 760

Db        782 VECFRIQD 789
```

RESULT 7
I54331
dipeptidyl aminopeptidase like protein - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I54331
R;Yokotani, N.; Doi, K.; Wenthold, R.J.; Wada, K.
Hum. Mol. Genet. 2, 1037-1039, 1993
A;Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-rela
A;Reference number: I54331; MUID:93372805; PMID:8103397
A;Accession: I54331
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-865 <RES>
A;Cross-references: UNIPROT:P42658; UNIPARC:UPI0000047378; GB:M96859; NID:g306705; PIDN
C;Superfamily: dipeptidyl-peptidase IV

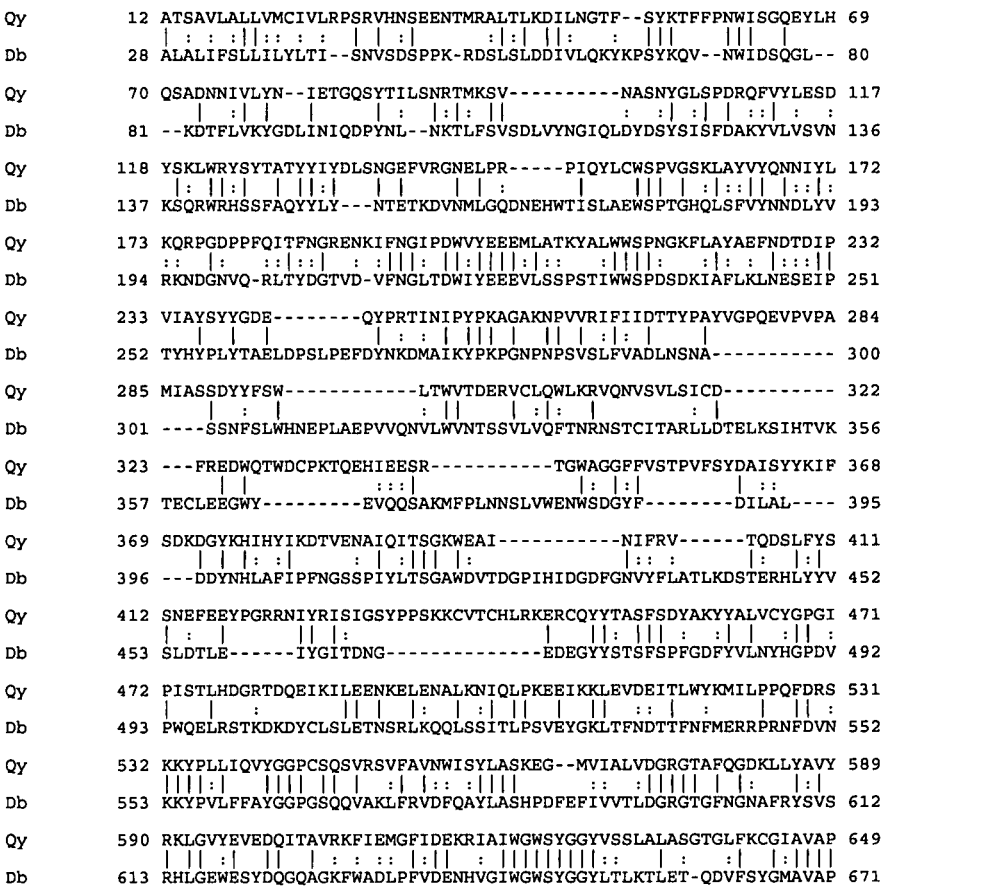
  Query Match   28.4%; Score 1161; DB 2; Length 865;
  Best Local Similarity   32.4%; Pred. No. 1.6e-72;
  Matches 255; Conservative 173; Mismatches 292; Indels   68; Gaps   23;

```
Qy         10 GVATSAVLALLVMC-------IVLRPSRVHNSEENTM---RALTLKDILNGTFSYKTFFP 59

Db         95 GIAI-ALLVILVICSLIVTSVILLTP-----AEDNSLSQKKKVTVEDLFSEDFKIHDPEA 148

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYS 119

Db        149 KWISDTEFIYREQKGTVRLWNVETNTSTVLIEGKKIESLRAIRYEISPDREYALFSYNVE 208
```



RESULT 8
T41703
dipeptidyl aminopeptidase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T41703
R;Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1998
A;Reference number: Z22011
A;Accession: T41703
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-793 <MUR>
A;Cross-references: UNIPROT:O14073; UNIPARC:UPI000013AB95; EMBL:AL031180; PIDN:CAA20138.
A;Experimental source: strain 972h-; cosmid c2E11 -chimeric
C;Genetics:
A;Gene: SPAC2E11.08
A;Map position: 1
C;Superfamily: dipeptidyl-peptidase IV

Query Match 21.2%; Score 864.5; DB 2; Length 793;
Best Local Similarity 29.1%; Pred. No. 6e-52;
Matches 238; Conservative 131; Mismatches 285; Indels 165; Gaps 29;



RESULT 9
S66261
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Flavobacterium meningosepticum
C;Species: Flavobacterium meningosepticum
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S66261
R;Kabashima, T.; Yoshida, T.; Ito, K.; Yoshimoto, T.
Arch. Biochem. Biophys. 320, 123-128, 1995
A;Title: Cloning, sequencing, and expression of the dipeptidyl peptidase IV gene from F
A;Reference number: S66261; MUID:95314307; PMID:7793970
A;Accession: S66261
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-711 <KAB>
A;Cross-references: UNIPROT:Q47900; UNIPARC:UPI00000B2115; EMBL:D42121; NID:g577283; PI

C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidylpeptide hydrolase

Query Match 20.6%; Score 841; DB 2; Length 711;
Best Local Similarity 29.7%; Pred. No. 2.2e-50;
Matches 232; Conservative 127; Mismatches 312; Indels 110; Gaps 24;

```
Qy          6 KIVFGVATSAVLA---LLVMCIVLRPSRVHNSEENTMRALTLKDI--LNGTFSYKTFFPN 60
              | :| : : ||:|   |    | |   :::: :    ||  :  |  ||   :| | |
Db          3 KKIFSLLSIAVVAFHGLSAQEITL--DKIYSGQ---YRAKGISGIASLNDGENYATIEPT 57

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              |:   |     : |||                :   |  | | |   : |:
Db         58 GIAKYSYKTSQKEKNIV------------------DGSFQGYTFSNDESKILLQKSSQS 98

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ::|:|:   : : || :   |  |      ||   :|| |||:|::  ||:: :
Db         99 IYRHSFLGKFEVKDLKSRTVVSLNN-ANWIQEPKFSPDGSKVAFIADNNLFYQDLNTGKI 157

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              ||| :|::|:| ||: |||||||      |   |:  |  | :  |::  :| |     |
Db        158 TQITTDGKKNEIINGLGDWVYEEEFGHADY-YQWNKAGDALVFVRFDERKVPEINIPIYY 216

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMI---ASSDYYFSWL- 296
              ||: :  |||||:|  |  ::    |   |      | :   :|  ||   |
Db        217 QNLYPKLMTYKYPKAGEENSAVTAYL----YQLSSGKS-----AQLNFGSSEKYYIPQLF 267

Qy        297 -TWVTDERVCLQWLKRVQNVSVLSICD--------FREDWQTW--DCPKTQEHIEESRTG 345
              |  || |     :   | :|  :          | |     |       | :::: 
Db        268 QTNANDEIVVATANRHQNKVDLLKVNTKTAAVSKLFTETDNAWIETDNLTMEFLDDNSFL 327

Qy        346 WAGGFFVSTPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQ 405
              ||                        |::||::|:::     :   |:: | || ||
Db        328 WA--------------------SERDGHRHLYWYDAAGKLKKQVSKGDWEIIN------ 360

Qy        406 DSLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYY-------TASFSDY 458
              :|  |     |  : :|  :       :|    ::   : |           :|:||
Db        361 ---YYGYN-----PKTKEVYIQTTEKGSINKVVSKLNINTGKTQLLSNAEGNNSAAFSKT 412

Qy        459 AKYYALVCYGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKEE---IKKLEVDE 515
              |:        :|   :      :::| |: | :| | ||:     :|   |      |:
Db        413 FNYFINTSSTAKVPTKYILKDANGKDVKELQNNDDLLNKLKSDNFIAKEFITIPNAAGDQ 472

Qy        516 ITLWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVN---WISYLASKEGMVIALV 572
              :  |  || |  || :||||: :  | || || | : :       |   || | | ::  |
Db        473 MNAW--MIKPKNFDPAKKYPVFMFQYSGPGSQQVANSWDGGNGIWFDMLAQK-GYLVVCV 529

Qy        573 DGRGTAFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSL 632
              ||||| |:| |     |: || ||:|||||| :      ::|: || |:|||||||:||
Db        530 DGRGTGFRGTKYKKVTYKNLGKYEIEDQITAAKWLGNQSYVDKSRIGIFGWSYGGYMASL 589

Qy        633 ALASGTGLFKCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNV 692
              |:  |  :|| ||||||::| :| |:|||||  | |  :| : |  ::    |:  :
Db        590 AMTKGADVFKMGIAVAPVTNWRFYDSIYTERF--LQTPQENKDGYDLNSPTTYAKLLKG- 646

Qy        693 DYLLIHGTADDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLSTN-HLYTHMTHF 751
               :|||||||||||||| : ::||:  :  |  | | |:|| : | :|   ||  ||::
Db        647 KFLLIHGTADDNVHFQNSMEFSEALIQNKKQFDFMAYPDKNHSIIGGNTRPQLYEKMTNY 706

Qy        752 L 752
              :
Db        707 I 707
```

RESULT 10
A30107
dipeptidyl aminopeptidase B (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YHR028c
C;Species: Saccharomyces cerevisiae
C;Date: 07-Jun-1990 #sequence_revision 30-May-1997 #text_change 09-Jul-2004
C;Accession: S46780; A30107
R;Du, Z.
submitted to the EMBL Data Library, June 1994

A;Description: The sequence of S. cerevisiae cosmid 8082.
A;Reference number: S46773
A;Accession: S46780
A;Molecule type: DNA
A;Residues: 1-818 <DUZ>
A;Cross-references: UNIPROT:P18962; UNIPARC:UPI0000031A5F; EMBL:U10399; NID:g500689; PI
R;Roberts, C.J.; Pohlig, G.; Rothman, J.H.; Stevens, T.H.
J. Cell Biol. 108, 1363-1373, 1989
A;Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an i
A;Reference number: A30107; MUID:89174971; PMID:2647766
A;Accession: A30107
A;Molecule type: DNA
A;Residues: 1-82,'H',84-124,'N',126-181,'LRRLET',189-199,'N',201-365,'DFKRGKERKF',376-5
A;Cross-references: UNIPARC:UPI0000172A3C; EMBL:X15484
A;Note: the authors translated the codon ACC for residue 572 as Asn
C;Genetics:
A;Gene: SGD:DAP2; STE13; MIPS:YHR028c
A;Cross-references: SGD:S0001070; MIPS:YHR028c
A;Map position: 8R
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein; yeast vac
F;30-45/Domain: transmembrane #status predicted <TMM>
F;63,79,110,139,392,421/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.9%; Score 814.5; DB 1; Length 818;
Best Local Similarity 27.2%; Pred. No. 1.9e-48;
Matches 229; Conservative 150; Mismatches 310; Indels 153; Gaps 30;

```
Qy         11 VATSAVLALLVM-CIVLRPSRVHNS------EENT------MRALTLKDILNGTFSYKTF 57
              :    :| ||:   ::|  |  |:|      | |:          ::   : | ||  |
Db         28 IRVGIILVLLIWGTVLLLKSIPHHSNTPDYQEPNSNYTNDGKLKVSFSVVRNNTFQPKYH 87

Qy         58 FPNWISGQEYLHQSADNNIVLYNIETGQSYTILS--------------------NRTMK 96
                |||     ::   |:: ||      || : |                    | |::
Db         88 ELQWISD----NKIESNDLGLYVTFMNDSYVVKSVYDDSYNSVLLEGKTFIHNGQNLTVE 143

Qy         97 SVNASNYGLSPDRQFVYLESDYSKLWRYSYTATYYIYDLSNGEFVR-GNELPRPIQYLCW 155
              |: |     ||| : : : ::  : ||:|   :|::|| |:  |   |||:   |    |
Db        144 SITA-----SPDLKRLLIRTNSVQNWRHSTFGSYFVYDKSSSSFEEIGNEVALAI----W 194

Qy        156 SPVGSKLAYVYQNNIYLKQRPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWS 215
              ||  : :|||  ||||:          :   |   : :||| ||||||||:     | |||
Db        195 SPNSNDIAYVQDNNIYIYSAISKKTIRAVTNDGSSFLFNGKPDWVYEEEVFEDDKAAWWS 254

Qy        216 PNGKFLAYAEFNDTDIPVIAYSYYGDEQ---YPRTINIPYPKAGAKNPVVRIFIID---- 268
              | | :||: : :::::       ||  ::   ||   :| |||:|  ||   :::
Db        255 PTGDYLAFLKIDESEVGEFIIPYYVQDEKDIYPEMRSIKYPKSGTPNPHAELWVYSMKDG 314

Qy        269 -TTYPAYVGPQEVPVPAMIASSDYYFSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDW 327
               : :|   | ::           : :||| :  | ::   |  ::  : : |
Db        315 TSFHPRISGNKK--------DGSLLITEVTWVGNGNVLVKTTDRSSDILTVFLID----- 361

Qy        328 QTWDCPKTQEHI--EESRTGW------------------AGGFFVSTPVFSYDAISYYK 366
                 ||   :  | |  ||                    |:   |:  |: ::|::
Db        362 ---TIAKTSNVVRNESSNGGWWEITHNTLFIPANETFDRPHNGYVDILPIGGYNHLAYFE 418

Qy        367 IFSDKDGYKHIHYIKDTVENAIQITSGKWEAIN---IFRVTQDSLFYSSNEFEEYPGRRN 423
               |:  ||              :| |||| :|    |   :: |:: |    :    |:
Db        419 -NSNSSHYK------------TLTEGKWEVVNGPLAFDSMENRLYFIST--RKSSTERH 462

Qy        424 IYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGRTDQ 483
              :| | : |  |::        ::    |  |||   ::  |   || :|   :  | : :
Db        463 VYYIDLRS--PNEIIEVTDTSED--GVYDVSFSSGRRFGLLTYKGPKVPYQKIVDFHSRK 518

Qy        484 EIKI------------LEENKELENALKNIQLPKEEIKKLEV-------DEITLWYKMIL 524
               |              ||:|: |   |::  :|::  ::| :       | :    |: ||
Db        519 AEKCDKGNVLGKSLYHLEKNEVLTKILEDYAVPRKSFRELNLGKDEFGKDILVNSYE-IL 577

Qy        525 PPQFDR--SKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGD 582
              |  ||   |  ||:    |||| || |   |:| :   :||:   :: :|||||| |:|
Db        578 PNDFDETLSDHYPVFFFAYGGPNSQQVVKTFSVGFNEVVASQLNAIVVVVDGRGTGFKGQ 637

Qy        583 KLLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGL-F 641
```

```
Db        638 DFRSLVRDRLGDYEARDQISAASLYGSLTFVDPQKISLFGWSYGGYLTLKTLEKDGGRHF 697

Qy        642 KCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEHY-----KNSTVMARAEYFRNVDYLL 696
Db        698 KYGMSVAPVTDWRFYDSVYTERYMHTP--QENFDGYVESSVHNVTALAQANRF-----LL 750

Qy        697 IHGTADDNVHFQNSAQIAKAL-VNAQVDFQAMWYSDQNHGLSGLSTNHL-YTHMTHFLKQ 754
Db        751 MHGTGDDNVHFQNSLKFLDLLDLNGVENYDVHVFPDSDHSIRYHNANVIVFDKLLDWAKR 810

Qy        755 CF 756
Db        811 AF 812


RESULT 11
T19514
hypothetical protein C27C12.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19514
R;Thomas, K.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19134
A;Accession: T19514
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-829 <WIL>
A;Cross-references: UNIPROT:Q18253; UNIPARC:UPI000007DC53; EMBL:Z69883; PIDN:CAA93743.1;
A;Experimental source: clone C27C12
C;Genetics:
A;Gene: CESP:C27C12.7
A;Map position: X
A;Introns: 4/3; 51/1; 123/3; 166/3; 188/2; 279/2; 392/3; 501/3; 553/1; 583/1; 606/2; 647
C;Superfamily: dipeptidyl-peptidase IV

  Query Match             19.0%;  Score 774.5;  DB 2;  Length 829;
  Best Local Similarity   28.3%;  Pred. No. 1.2e-45;
  Matches  241;  Conservative  140;  Mismatches  303;  Indels  167;  Gaps   37;

Qy         16 VLALLVMCIVLRPSRVHNSEEN-----------------TMRALTLKDILNG--TFSYK 55
Db         29 VVFILIICVIFVFSVIFTFMQNPINLNSDNGFNQTSGNTSSLEATTLKPKFSSLMTTTRR 88

Qy         56 TFFPNWISGQEYLHQSADNNIVLYN---------IETGQSYTILSNRTMKSV-------- 98
Db         89 FTFEQLFSGKQFL-------VDYYDYIWLPDGSFVQMNDDFTI--RKQMKKIPLGSSVAE 139

Qy         99 ---NASNY--GLSPDRQFVYLESDYSKLWRYSYTATYYIYDLSN-----GEFVRGNELPR 148
Db        140 PFFNNGEYVKALSSNMKYAYGSKKVNELWRHSAEYLYHIVKINNKTVSTEQWHVGPEENS 199

Qy        149 PIQYLCWSPVGSK--LAYVYQNNIYLKQRPGDP--PFQITFNGRENKIFN--GIPDWVYE 202
Db        200 LIQAFYWNPNASSNDFVYVHNYNLYYQKDPEKPDGAIQLTVGG---STFNRFGLANWLYE 256

Qy        203 EEMLATKYALWWSPNGKFLAYAEFNDTDI-PVIAYSYYGDEQYPRTINIPYPKAGAK-NP 260
Db        257 EEILEASSAVWWSPSGRYVSYLRFDDREVNRIFLPKYTDDDSYVEYFELPYPKAGVQNNT 316

Qy        261 VVRIFIIDT-------TYPAYVGPQEVPVPAMIASSDYYFSWLTWVT---------DERV 304
Db        317 LVTQYIWDSENHKIVETAP----PNELSA----ANGDYYVLTNKWITMPRNGSDLGEERL 368

Qy        305 CLQWLKRVQNVSVLSICDFR------------EDWQTWDCPKTQEHIEESRTGWAGGFFV 352
Db        369 VTVWANRDQNHVYFSLCNEQDCVMALSFQFSIDNRQLWVSPKDVRGVFPTETG-----FL 423

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINI-------FRVTQ 405
Db        424 TVLPHKHD---------DGNIYNHVAHVELDGTGTGKIT--KWIGENFDVILVLGYSSKI 472

Qy        406 DSLFYSS-----NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRK-----ERCQYYTASF 455
Db        473 DALTFSAYGDGVGEFSTYIVREAMY---------SNKKTT--LQKVTDQFEDCKTLGSQS 521

Qy        456 SDYAKYYALV-CYGPGIPISTLHDGRTDQEIKILEENKE--LENALKNI---QLPKEEIK 509
Db        522 ADPTGQRIVVQCEKP-------FDNTRLYLVDVVDTTKKIMLEGGTKAVIPFDVPNMKFG 574

Qy        510 KLEVDE-ITLWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMV 568
Db        575 KLKLPSGIDGHYMMLTPANLLDGAKIPLLLDIYGGPDSKQVFQKTPTAHAIQIVSQYDIA 634

Qy        569 IALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQITAVRKFIE-MGFIDEKRIAIWGWSYGG 627
Db        635 YARIDVRGTGGRGWDVKEAVYRKLGDAEVVDTLDMIRAFINTFGFIDEDRIAVMGWSYGG 694

Qy        628 YVSS-LALASGTGLFKCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARA 686
Db        695 FLTSKIAIKDQGELVKCAISIAPVTDFKYYDSAYTERYLGQPA--ENLQGYINTNVIPHA 752

Qy        687 EYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLSTNHLYT 746
Db        753 RNVTNVKYLLAHGERDDNVHYQNSARWSEALQQNGIHFTQLVYANEAHSLSH-KLFHLYG 811

Qy        747 HMTHFL-KQCF 756
Db        812 EVQRFLMNDCF 822


RESULT 12
T25173
hypothetical protein T23F1.7a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25173
R;Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19990
A;Accession: T25173
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-779 <WIL>
A;Cross-references: UNIPROT:O18119; UNIPARC:UPI000002A220; EMBL:Z81129; PIDN:CAB03411.1,
A;Experimental source: clone T23F1
C;Genetics:
A;Gene: CESP:T23F1.7a
A;Map position: 5
A;Introns: 13/3; 52/1; 111/3; 151/2; 177/3; 346/3; 504/1; 537/1; 628/2; 683/1; 723/1
C;Superfamily: dipeptidyl-peptidase IV

  Query Match             18.2%;  Score 744;  DB 2;  Length 779;
  Best Local Similarity   27.9%;  Pred. No. 1.4e-43;
  Matches  225;  Conservative  137;  Mismatches  327;  Indels  118;  Gaps   30;

Qy         10 GVATSAVLALLVMCIVLRP----SRVHNSEENTMRALTLKDILNGTFSYKTF-------- 57
Db         27 GITTAIVVVLLILVMIFAALVFFTPLFAAKSFGSWRLNVSDLRSLRYPYAEFAFTDNNAV 86

Qy         58 -FPNWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSV--NASNYGLSPDRQFVYL 114
Db         87 VMQSW-EGVEIVEDGVSR--LIFGRENGAEITPSADRKYFAMMDHAPNPGMNPQNETFHL 143

Qy        115 ESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQ 174
Db        144 -----KIVNNNERITYDI-GLRKEESV--------IQAFKWNGKFNDFVFVESNKIYYQS 189

Qy        175 RPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEF-NDTDIPV 233
Db        190 SPEEEGLTRVSNGGEHTV-DGLFDWIYEEEIFGRKDAMWWSTKGDQLAYASYDNHLTKNV 248

Qy        234 IAYSYYGDEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYF 293
Db        249 SLKTYHRLEPYPIDTNFHYPKTFAK--VLPTYTLSIWNKKTEQSRQLDVQLKDSLSYHYL 306
```

```
Qy        294 ---SWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGF 350
Db        307 LAVKWLEINGTEQLVSVWTNRYQNEVALTIC----DWDTAIC-----RLE---------- 347

Qy        351 FVSTPVFSYDAISYYKIFSDKDGYKHIHYIKDTV---------ENAIQ------------ 389
Db        348 ------FEYKYAS--KRWVTHDDFHSITSFEDTLFFLLPHDKRDNAFQQVASLRLSHGQL 399

Qy        390 -----ITSGKWEAINIFRVTQDSLFYSSNEFEEYPGRRNIYRISIGSYP-PSKKCVTCHL 443
Db        400 RTPKFLNLGEYDVTSINGINKETRTIFFHAAAPKPSHRSLFSYSLADESRNSAYCISCSI 459

Qy        444 RKERCQYYTASFSDYAKYYALVCYGPGIPISTLHD-GRTDQEIKILEEN----KELENAL 498
Db        460 K--NCTWAQAQMDDQMKTAIVSCKGPAAPHTAIVNLTRMDSDKKTEHANLLYDKTYQNRV 517

Qy        499 KNIQLP---KEEIKKLEVDEITLWYKMILPPQF---DRSKKYPLLIQVYGGPCSQSVRSV 552
Db        518 EEAGLPVIIKETIK--ISDDFDALIKLSIPKDIYNRDKHQAIPLIVHVYGGPNDQNTKEA 575

Qy        553 FAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEM-- 610
Db        576 TQIGIEEVVASASQAAILRIDGRGSGGRGWKYRSAIYGQLGTVEVEDQIKAIKVVLRLYR 635

Qy        611 GFIDEKRIAIWGWSYGGYVS-SLALASGTGLFKCGIAVAPVSSWEYYASVYTERFMGLPT 669
Db        636 HLLDARRVAVFGWSYGGFMTLSMVNEAPEQFFKCAVSVAPVTNFAYYDATYTERYMG--- 692

Qy        670 KDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNAQVDFQAMWY 729
Db        693 -DAPLESY--SDVTKKLDNFKSTRLLLMHGLLDDNVHFQNSAILIDELQNRGVDFDLMVY 749

Qy        730 SDQNHGLSGLSTNHLYTHMTHFLKQCF 756
Db        750 PNQAHSLSS-RTSHVVGKMTHFLRQCF 775
```

RESULT 13
T25174
hypothetical protein T23F1.7b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25174
R;Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19990
A;Accession: T25174
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-799 <WIL>
A;Cross-references: UNIPROT:O18119; UNIPARC:UPI0000061190; EMBL:Z81129; PIDN:CAB03412.1;
A;Experimental source: clone T23F1
C;Genetics:
A;Gene: CESP:T23F1.7b
A;Map position: 5
A;Introns: 13/3; 52/1; 111/3; 197/3; 366/3; 524/1; 557/1; 648/2; 703/1; 743/1
C;Superfamily: dipeptidyl-peptidase IV

  Query Match 18.1%; Score 739; DB 2; Length 799;
  Best Local Similarity 27.3%; Pred. No. 3.2e-43;
  Matches 224; Conservative 143; Mismatches 329; Indels 124; Gaps 31;

```
Qy         10 GVATSAVLALLVMCIVLRP----SRVHNSEENTMRALTLKDILNGTFSYKTF-------- 57
Db         27 GITTAIVVVLLILVMIFAALVFFTPLFAAKSFGSWRLNVSDLRSLRYPYAEFAFTDNNAV 86

Qy         58 -FPNWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSV--NASNYGLSPDRQFVYL 114
Db         87 VMQSW-EGVEIVEDGVSR--LIFGRENGAEITPSADRKYFAMMDHAPNPGMNPQNETFHL 143

Qy        115 E------------SDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSK 161
Db        144 KIVNNNERLNPLLPFEVEELFR-ELSDSRITYDIG----LRKEE--SVIQAFKWNGKFND 196

Qy        162 LAYVYQNNIYLKQRPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFL 221
Db        197 FVFVESNKIYYQSSPEEEGLTRVSNGGEHTV-DGLPDWIYEEEIFGRKDAMWWSTKGDQL 255

Qy        222 AYAEF-NDTDIPVIAYSYYGDEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEV 280
Db        256 AYASYDNHLTKNVSLKTYHRLEPYPIDTNFHYPKTFAK--VLPTYTLSIWNKKTEQSRQL 313

Qy        281 PVPAMIASSDYYF---SWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQE 337
Db        314 DVQLKDSLSYHYLLAVKWLEINGTEQLVSVWTNRYQNEVALTIC----DWDTAIC----- 364

Qy        338 HIEESRTGWAGGFFVSTPVFSYDAISYYKIFSDKDGYKHIHYIKDTV---------ENAI 388
Db        365 RLE----------------FEYKYAS--KRWVTHDDFHSITSFEDTLFFLLPHDKRDNAF 406

Qy        389 Q-----------------ITSGKWEAINIFRVTQDSLFYSSNEFEEYPGRRNIYRISIGS 431
Db        407 QQVASLRLSHGQLRTPKFLNLGEYDVTSINGINKETRTIFFHAAAPKPSHRSLFSYSLAD 466

Qy        432 YP-PSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHD-GRTDQEIKILE 489
Db        467 ESRNSAYCISCSIK--NCTWAQAQMDDQMKTAIVSCKGPAAPHTAIVNLTRMDSDKKTEH 524

Qy        490 EN----KELENALKNIQLP---KEEIKKLEVDEITLWYKMILPPQF---DRSKKYPLLIQ 539
Db        525 ANLLYDKTYQNRVEEAGLPVIIKETIK--ISDDFDALIKLSIPKDIYNRDKHQAIPLIVH 582

Qy        540 VYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVED 599
Db        583 VYGGPNDQNTKEATQIGIEEVVASASQAAILRIDGRGSGGRGWKYRSAIYGQLGTVEVED 642

Qy        600 QITAVRKFIEM--GFIDEKRIAIWGWSYGGYVS-SLALASGTGLFKCGIAVAPVSSWEYY 656
Db        643 QIKAIKVVLRLYRHLLDARRVAVFGWSYGGFMTLSMVNEAPEQFFKCAVSVAPVTNFAYY 702

Qy        657 ASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKA 716
Db        703 DATYTERYMG----DAPLESY--SDVTKKLDNFKSTRLLLMHGLLDDNVHFQNSAILIDE 756

Qy        717 LVNAQVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCF 756
Db        757 LQNRGVDFDLMVYPNQAHSLSS-RTSHVVGKMTHFLRQCF 795
```

RESULT 14
A49737
dipeptidyl aminopeptidase (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein O5045; protein YOR219c; protein YOR50-9
C;Species: Saccharomyces cerevisiae
C;Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 09-Jul-2004
C;Accession: A49737; S45451; S60946; S67112; S71721
R;Santa Anna-A, S.; Herskowitz, I.
submitted to the Protein Sequence Database, July 1993
A;Reference number: A49737
A;Accession: A49737
A;Molecule type: DNA
A;Residues: 1-931 <SAN>
A;Cross-references: UNIPROT:P33894; UNIPARC:UPI0000136060; GB:L21944; NID:g347196; PIDN
R;Anna-Arriola, S.S.; Herskowitz, I.
Yeast 10, 801-810, 1994
A;Title: Isolation and DNA sequence of the STE13 gene encoding dipeptidyl aminopeptidas
A;Reference number: S45451; MUID:95066382; PMID:7975897
A;Accession: S45451
A;Molecule type: DNA
A;Residues: 1-931 <ANN>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:L21944; NID:g347196; PIDN:AAA35119.1; P
R;Galisson, F.; Dujon, B.
submitted to the EMBL Data Library, October 1995
A;Description: Sequence and analysis of a 33 kb fragment from the right arm of chromoso
A;Reference number: S60938

```
A;Accession: S60946
A;Molecule type: DNA
A;Residues: 1-931 <GAL>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1; P
R;Boyer, J.; Fairhead, C.; Gaillon, L.; Galisson, F.; Michaux, G.; Thierry, A.; Dujon, B
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67104
A;Accession: S67112
A;Molecule type: DNA
A;Residues: 1-931 <BOY>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:Z75127; NID:g1420507; PIDN:CAA99437.1; P
A;Experimental source: strain S288C
R;Galisson, F.; Dujon, B.
Yeast 12, 877-885, 1996
A;Title: Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV c
A;Reference number: S71713; MUID:96437977; PMID:8840505
A;Accession: S71721
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-931 <GAW>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1; P
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C;Genetics:
A;Gene: SGD:STE13; YCI1
A;Cross-references: SGD:S0005745; MIPS:YOR219c
A;Map position: 15R
C;Function:
A;Description: involved in processing of alpha-factor prepropheromone
C;Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein
F;113-150/Domain: transmembrane #status predicted <TMM>
F;377/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;785,863,896/Active site: Ser, Asp, His #status predicted

  Query Match             17.6%;  Score 721;  DB 2;  Length 931;
  Best Local Similarity   25.4%;  Pred. No. 7.2e-42;
  Matches  215;  Conservative  152;  Mismatches  336;  Indels  144;  Gaps   28;

Qy          7 IVFGVATSAVLALLVMCI--------VLRPSRVHNSEENTMRALTLKDILNGTFSYKTFF 58
              :|| :   :|| |||: |        : ||          |: :::::||| |:
Db        122 LVFTLIALSVLVLLVILIPSKLLPTKITRPKTSAGDSSLGKRSFSIENVLNGDFAIPEDT 181

Qy         59 PNWIS-GQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVY---- 113
               ::|   |  | | :|  :     |       :: :        | |     :|:|
Db        182 FHFIDPPQRLLGQDSDPGLYFTTKEIDGHTNFIAKQLFDETFEVNLG---GNRFLYEGVE 238

Qy        114 --------------------LESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPI--- 150
                                 |||::    |:|   |:| ||: |          ||
Db        239 FTVSTVQINYKLDKLIFGTNLESEF----RHSSKGFYWIKDLNTGNI-------EPILPP 287

Qy        151 --------------QYLCWSPVGSKLAYVYQNNIYLKQRPGDPPFQITFNGRENKIFNGI 196
                           |  :||  : : :||:||::|:|       ::| :| :: |||
Db        288 EKSDDNYELGLSKLSYAHFSPAYNYIYFVYENNLFLQQVNSGVAKKVTEDGSKD-IFNAK 346

Qy        197 PDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG--DEQYPRTINIPYPK 254
              |||:||||:||:  |:||:|:     :| |||| :  |  : |   :| |      | |||
Db        347 PDWIYEEEVLASDQAIWWAPDDSKAVFARFNDTSVDDIRLNRYTNMNEAYLSDTKIKYPK 406

Qy        255 AGAKNPVVRIFIID----TTYPAYVGPQEVPVPAMIASSDYYFSWLT------WVTDERV 304
              | :||   :|:::     |   | |:       | |   |::       :||
Db        407 PGFQNPQFDLFLVNLQNGIIYSINTGGQK-------DSILYNGKWISPDTFRFEITD--- 456

Qy        305 CLQWLKRVQNVSVLSICDFREDWQTWDCPKTQE-HIEESRTGWAGGFFVST--------- 354
                    :|  :|       | : :| | :|   :  : :    |:  |
Db        457 --------RNSKIL-------DVKVYDIPSSQMLTVRNTNSNLFNGWIEKTKDILSIPPK 501

Qy        355 PVFSYDAISYYKIFSDKDGYKHIHYIKDT-VENAIQITSGKWE----AINIFRVTQDSLF 409
              |       | | :|  |: |: |      :  ||:| | ||     | :    |::|
Db        502 PELKRMDYGYIDIHADSRGFSHLFYYPTVFAKEPIQLTKGNWEVTGNGIVGYEYETDTIF 561

Qy        410 YSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGP 469
              :::||     :::| ||:      :      :: :|   |  |:|      ||
Db        562 FTANEIGVM--SQHLYSISLTDSTTQNTFQSLQNPSDKYDFYDFELSSSARYAISKKLGP 619

Qy        470 GIPI-----------STLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDE-I 516
                ||           : :|    |  |  | :::: :  :|| ||    | : :|: :
Db        620 DTPIKVAGPLTRVLNVAEIH----DDSILQLTKDEKFKEKIKNYDLPITSYKTMVLDDGV 675

Qy        517 TLWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRG 576
              : |  | |   :  ||||:|: :|||| ||:  :  :: :   : |   :::  :: ||
Db        676 EINYIEIKPANLNPKKKYPILVNIYGGPGSQTFTTKSSLAFEQAVVSGLDVIVLQIEPRG 735

Qy        577 TAFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEMG--FIDEKRIAIWGWSYGGYVSSLAL 634
              |  :|         ||| :|  |      :|||:     ||| :||||||||: |   :
Db        736 TGGKGWSFRSWAREKLGYWEPRDITEVTKKFIQRNSQHIDESKIAIWGWSYGGFTSLKTV 795

Qy        635 ASGTG-LFKCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVD 693
                  |  ||  :|||||::|  | ||||||:|  |:  :| | |  : :   : | ::
Db        796 ELDNGDTFKYAMAVAPVTNWTLYDSVYTERYMNQPS--ENHEGYFEVSTIQNFKSFESLK 853

Qy        694 YL-LIHGTADDNVHFQNSAQIAKAL-VNAQVDFQAMWYSDQNHGLSGLSTNHL-YTHMTH 750
              | ::||| ||||| ||: ::   | :    ::    : | :| :   :   : : :
Db        854 RLFIVHGTFDDNVHIQNTFRLVDQLNLLGLTNYDMHIFPDSDHSIRYHNAQRIVFQKLYY 913

Qy        751 FLKQCFS 757
              :|:  |:
Db        914 WLRDAFA 920


RESULT 15
A87516
dipeptidyl peptidase IV [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: A87516
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87516
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-738 <STO>
A;Cross-references: UNIPROT:Q9A6E0; UNIPARC:UPI00000C7616; GB:AE005673; NID:g13423647;
C;Genetics:
A;Gene: CC2154

  Query Match             15.2%;  Score 622.5;  DB 2;  Length 738;
  Best Local Similarity   25.8%;  Pred. No. 3.6e-35;
  Matches  207;  Conservative  124;  Mismatches  326;  Indels  145;  Gaps   27;

Qy         12 ATSAVLALLVMCIV--LRPSRVHNSEE---NTMRALTLKDILNGTFSYKTFFPNWISGQE 66
              | :|||||     :   | | |:         | | : |
Db          8 ALAAVLALSTPALAQKLTPERIFGDPSLSGPTARGVAL---------------------- 45

Qy         67 YLHQSADNNIVLY----------------NIETGQSYTILSNRTMKS------------- 97
                  | |   | |                   ::: |: | :: :  : |
Db         46 ----SPDGKRVTYLKGKPEAANIQDLWAADVKGGEPYRLIDSAALSSGDKELSEAEKARR 101

Qy         98 --VNASNYGL---SPDRQ--FVYLESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPI 150
                   |  |:    | |||  |: :  |                 |:  :::|:  |  | |
Db        102 ERARVSARGIVEYSWDRQGRFILVPLD----------GDLYLDAVADGKITRLTETPGDE 151

Qy        151 QYLCWSPVGSKLAYVYQNNIYLKQRPGDPPFQITFNGRENKIFNGIPDWVYEEEM-LATK 209
                  || |  ::||    |:|:|    |      :| :|::    | |: ::: ::||:   |
Db        152 VDAKVSPKGGYVSYVRDQNLYIKPVAGGAETALTTDGKDALSF-GVAEFIVQEELDRFTG 210

Qy        210 YALWWSPNGKFLAYAEFNDTDIPVIAYSYYGDEQYPRTINIPYPKAGAKNPVVRIFIIDT 269
              |  ||||:   : |   ::: : ::  :  |        :|  ||:||  | || :|: |
Db        211 Y--WWSPDESRIVYTRVDESGVDIVPRADIGPGG-ATVVNQRYPRAGRPNAVVDLFVRD- 266

Qy        270 TYPAYVGPQEVPVPAMIASSDYYFSWLTWVTD-ERVCLQWLKRVQNVSVLSICDFREDWQ 328
                    :  :|    : |: | | : : |  | : | :| | | |    | :  |
Db        267 -----LASGKVTALDLGANKDIYVARVAWSADGKTVYVQRLSRDQ--KTLDLLAF----- 314
```

```
Qy        329 TWDCPKTQEHIEESRTGWAGGFFVSTPVFSYDAISYYKIFSD--------KDGYKHIHYI 380
Db        315 ------------DAATGAGKTILTDTDPHFIEVSNDFRPLTDGTFLWGSEKDGNQHLYRY 362

Qy        381 KDTVENAIQITSGKWEAINIFRVTQD---SLFYSSNEFEEYPGRRNIYRISIGSYPPSKK 437
Db        363 AADGKLIAQITKGDWPVIGLEGVDEARKVAIFSAS---IDTPIERRLYEVSYAK-PGKPK 418

Qy        438 CVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGRTDQEIKILEENKELENA 497
Db        419 ALT-----SAGGWWAAKVADNGGAFAGTYSDPKTPSQTALYSADGKRVRWIEENKLAEGH 473

Qy        498 ---LKNIQLPKEEIKKLE-VDEITLWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVF 553
Db        474 PYWPYAANLPQPEFGSLKAADGETLHYEILKPIGFDPAKKYPAIVSVYGGPHAQRVMK-- 531

Qy        554 AVNWIS-----YLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQITAVRKFI 608
Db        532 --NWHSPSERTYL--EAGYVIFKLDNRGSGNRSAKFMRALDRKLGTVEVEDQLLGAKFLA 587

Qy        609 EMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVYTERFMGLP 668
Db        588 SQPYVDADKLGVMGWSYGGFMALMLLTAENTPFKAGAAGAPPTEWSLYDTAYTERYMGKP 647

Qy        669 TKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNAQVDFQAMW 728
Db        648 --DENKAGYAYSDINNRIDKLAPGSLLLLHGMADDNVIFENSTRLMAALQRKAILFEMAM 705

Qy        729 YSDQNHGLSGLSTNHLYTHMTH 750
Db        706 YPGERHSAPGSKTKGLSVLKTH 727
```

Search completed: January 28, 2006, 00:35:17
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2006, 00:25:20 ; Search time 160 Seconds
(without alignments)
3351.259 Million cell updates/sec

Title: US-10-681-659-1
Perfect score: 4086
Sequence: 1 MKTWVKIVFGVATSAVLALL.....TNHLYTHMTHFLKQCFSLSD 760

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4086	100.0	760	2	Q53TP5_HUMAN	Q53tp5 homo sapien
2	4069	99.6	760	1	SEPR_HUMAN	Q12884 homo sapien
3	3705.5	90.7	761	1	SEPR_MOUSE	P97321 mus musculu
4	3686.5	90.2	761	2	Q8R492_RAT	Q8r492 rattus norv
5	2485.5	60.8	755	2	Q91651_XENLA	Q91651 xenopus lae
6	2324	56.9	751	2	Q75S82_AGKHB	Q75s82 agkistrodon
7	2324	56.9	751	2	Q75S83_AGKHB	Q75s83 agkistrodon
8	2276	55.7	759	2	Q5ZI81_CHICK	Q5zi81 gallus gall
9	2208.5	54.1	766	1	DPP4_PIG	P22411 s dipeptidy
10	2203	53.9	767	1	DPP4_RAT	P14740 r dipeptidy
11	2190.5	53.6	766	1	DPP4_HUMAN	P27487 h dipeptidy
12	2190.5	53.6	766	2	Q53TN1_HUMAN	Q53tn1 homo sapien
13	2189.5	53.6	748	2	P70092_XENLA	P70092 xenopus lae
14	2189	53.6	765	2	Q5R7G7_PONPY	Q5r7g7 pongo pygma
15	2183.5	53.4	760	1	DPP4_MOUSE	P28843 m dipeptidy
16	2165.5	53.0	737	2	Q641D6_XENLA	Q641d6 xenopus lae
17	2165	53.0	765	1	DPP4_FELCA	Q9n2i7 f dipeptidy
18	2160	52.9	765	1	DPP4_BOVIN	P81425 b dipeptidy
19	2001.5	49.0	860	2	Q4RQJ4_TETNG	Q4rqj4 tetraodon n
20	1743	42.7	704	2	Q4S309_TETNG	Q4s309 tetraodon n
21	1668	40.8	350	2	Q6P7D6_RAT	Q6p7d6 rattus norv
22	1261	30.9	731	2	Q7PSF9_ANOGA	Q7psf9 anopheles g
23	1226.5	30.0	796	1	DPP10_HUMAN	Q8n608 homo sapien
24	1213.5	29.7	796	1	DPP10_RAT	Q6q629 rattus norv
25	1212	29.7	796	2	Q6INB7_XENLA	Q6inb7 xenopus lae
26	1192.5	29.2	797	1	DPP10_MOUSE	Q6nxk7 mus musculu
27	1177	28.8	863	1	DPP6_BOVIN	P42659 bos taurus
28	1161	28.4	865	1	DPP6_HUMAN	P42658 homo sapien
29	1159	28.4	803	1	DPP6_PANTR	Q5is50 pan troglod
30	1141.5	27.9	802	2	Q9VUH1_DROME	Q9vuh1 drosophila
31	1140	27.9	803	2	Q9QVT8_9MURI	Q9qvt8 rattus sp.
32	1137	27.8	859	1	DPP6_RAT	P46101 rattus norv
33	1132	27.7	801	2	Q5U4C2_MOUSE	Q5u4c2 mus musculu
34	1132	27.7	803	2	Q80VM5_MOUSE	Q80vm5 mus musculu
35	1132	27.7	804	1	DPP6_MOUSE	Q9z218 mus musculu
36	1109	27.1	724	2	Q5IS95_9PRIM	Q5is95 saimiri bol
37	1093	26.7	935	2	Q9VMB4_DROME	Q9vmb4 drosophila
38	1080.5	26.4	815	2	Q4RK63_TETNG	Q4rk63 tetraodon n
39	1066	26.1	439	2	Q6PG25_XENLA	Q6pg25 xenopus lae
40	1046.5	25.6	745	2	Q9VMM2_DROME	Q9vmm2 drosophila
41	1038	25.4	801	2	Q4RK66_TETNG	Q4rk66 tetraodon n
42	1018.5	24.9	604	2	Q53SS4_HUMAN	Q53ss4 homo sapien
43	986	24.1	842	2	Q4RVF8_TETNG	Q4rvf8 tetraodon n
44	976	23.9	775	2	Q5J6J3_TRIRU	Q5j6j3 trichophyto
45	960.5	23.5	895	2	Q7SHU8_NEUCR	Q7shu8 neurospora

ALIGNMENTS

RESULT 1
Q53TP5_HUMAN
ID Q53TP5_HUMAN PRELIMINARY; PRT; 760 AA.
AC Q53TP5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein FAP.
GN Name=FAP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Cotton M., Maupin R., Hawkins M., Harkins R.;
RT "The sequence of Homo sapiens BAC clone RP11-576I16.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AC007750; AAY24205.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Hypothetical protein; Transmembrane.
SQ SEQUENCE 760 AA; 87713 MW; 7FF817B5A4F75142 CRC64;

Query Match 100.0%; Score 4086; DB 2; Length 760;
Best Local Similarity 100.0%; Pred. No. 6.3e-281;
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

```
Qy          61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy         121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy         181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240

Qy         241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy         301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360

Qy         361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy         421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy         481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy         541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy         601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy         661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy         721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
               ||||||||||||||||||||||||||||||||||||||||
Db         721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
```

RESULT 2
SEPR_HUMAN
ID SEPR_HUMAN STANDARD; PRT; 760 AA.
AC Q12884; O00199; Q86Z29; Q99998; Q9UID4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Seprase (EC 3.4.21.-) (Fibroblast activation protein alpha) (Integral
DE membrane serine protease) (170-kDa melanoma membrane-bound
DE gelatinase).
GN Name=FAP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Fibroblast;
RX MEDLINE=94261645; PubMed=7911242;
RA Scanlan M.J., Raj B.K.M., Calvo B., Garin-Chesa P., Sanz-Moncasi M.P.,
RA Healey J.H., Old L.J., Rettig W.J.;
RT "Molecular cloning of fibroblast activation protein alpha, a member of
RT the serine protease family selectively expressed in stromal
RT fibroblasts of epithelial cancers.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:5657-5661(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND VARIANT THR-354.
RC TISSUE=Melanoma;
RX MEDLINE=97388251; PubMed=9247085; DOI=10.1016/S0925-4439(97)00032-X;
RA Goldstein L.A., Ghersi G., Pineiro-Sanchez M.L., Salamone M., Yeh Y.,
RA Flessate D., Chen W.-T.;
RT "Molecular cloning of seprase: a serine integral membrane protease
RT from human melanoma.";
RL Biochim. Biophys. Acta 1361:11-19(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND PROTEIN SEQUENCE OF 220-229;
RP 461-472 AND 511-518.
RC TISSUE=Melanoma;
RX MEDLINE=97218181; PubMed=9065413; DOI=10.1074/jbc.272.12.7595;
RA Pineiro-Sanchez M.L., Goldstein L.A., Dodt J., Howard L., Yeh Y.,
RA Chen W.-T.;
RT "Identification of the 170-kDa melanoma membrane-bound gelatinase
RT (seprase) as a serine integral membrane protease.";
RL J. Biol. Chem. 272:7595-7601(1997).
RN [4]
RP NUCLEOTIDE SEQUENCE (ISOFORM 2).
RC TISSUE=Melanoma;
RX MEDLINE=20112818; PubMed=10644713; DOI=10.1074/jbc.275.4.2554;
RA Goldstein L.A., Chen W.-T.;
RT "Identification of an alternatively spliced seprase mRNA that encodes
RT a novel intracellular isoform.";
RL J. Biol. Chem. 275:2554-2559(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1), AND VARIANT
RP THR-354.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP PROTEIN SEQUENCE OF 192-208; 220-240 AND 510-521.
RX MEDLINE=94327249; PubMed=7519584;
RA Rettig W.J., Su S.L., Fortunato S.R., Scanlan M.J., Raj B.K.M.,
RA Garin-Chesa P., Healey J.H., Old L.J.;
RT "Fibroblast activation protein: purification, epitope mapping and
RT induction by growth factors.";
RL Int. J. Cancer 58:385-392(1994).
CC -!- FUNCTION: May have a role in tissue remodeling during development
CC and wound healing, and may contribute to invasiveness in malignant
CC cancers.
CC -!- CATALYTIC ACTIVITY: Degrades gelatin and heat-denatured type I and
CC type IV collagen, but not native type I or type IV collagen. Does
CC not cleave laminin, fibronectin, fibrin or casein.
CC -!- SUBUNIT: Homodimer, or heterodimer with DPP4. The monomer is
CC inactive.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Found in cell
CC surface lamellipodia, invadopodia and on shed vesicles.

```
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1; Synonyms=L;
CC         IsoId=Q12884-1; Sequence=Displayed;
CC         Note=Major isoform;
CC       Name=2; Synonyms=S, Truncated;
CC         IsoId=Q12884-2; Sequence=VSP_005367;
CC   -!- TISSUE SPECIFICITY: Fibroblast specific.
CC   -!- INDUCTION: In fibroblasts at times and sites of tissue remodeling
CC       during development, tissue repair, and carcinogenesis.
CC   -!- PTM: N-glycosylated.
CC   -!- PTM: The N-terminus may be blocked.
CC   -!- SIMILARITY: Belongs to the peptidase S9B family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; U09278; AAB49652.1; -; mRNA.
DR   EMBL; U76833; AAC51668.1; -; mRNA.
DR   EMBL; AF007822; AAF21600.1; -; mRNA.
DR   EMBL; BC026250; AAH26250.1; -; mRNA.
DR   PDB; 1Z68; X-ray; A=39-757, B=-.
DR   MEROPS; S09.007; -.
DR   Ensembl; ENSG00000078098; Homo sapiens.
DR   HGNC; HGNC:3590; FAP.
DR   MIM; 600403; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0030027; C:lamellipodium; IDA.
DR   GO; GO:0005886; C:plasma membrane; NAS.
DR   GO; GO:0004274; F:dipeptidyl-peptidase IV activity; NAS.
DR   GO; GO:0004222; F:metalloendopeptidase activity; TAS.
DR   GO; GO:0042803; F:protein homodimerization activity; NAS.
DR   InterPro; IPR002471; Pept_S9_AS.
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR002469; Peptidase_S9B.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF00930; DPPIV_N; 1.
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW   3D-structure; Alternative splicing; Direct protein sequencing;
KW   Glycoprotein; Hydrolase; Polymorphism; Protease; Serine protease;
KW   Signal-anchor; Transmembrane.
FT   TOPO_DOM      1      4       Cytoplasmic (Potential).
FT   TRANSMEM      5     25       Signal-anchor for type II membrane
FT                                protein (Potential).
FT   TOPO_DOM     26    760       Extracellular (Potential).
FT   ACT_SITE    624    624       Charge relay system (By similarity).
FT   ACT_SITE    702    702       Charge relay system (By similarity).
FT   ACT_SITE    734    734       Charge relay system (By similarity).
FT   CARBOHYD     49     49       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     92     92       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     99     99       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    314    314       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    679    679       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC      1    521       Missing (in isoform 2).
FT                                /FTId=VSP_005367.
FT   VARIANT     354    354       R -> T (in dbSNP:1126507).
FT                                /FTId=VAR_019691.
FT   CONFLICT    207    207       P -> A (in Ref. 2 and 5).
FT   CONFLICT    229    229       K -> T (in Ref. 2 and 5).
SQ   SEQUENCE   760 AA;  87821 MW;  A0D34B4801BE07EA CRC64;

  Query Match             99.6%;  Score 4069;  DB 1;  Length 760;
  Best Local Similarity   99.6%;  Pred. No. 1e-279;
  Matches  757;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              |||||||||||||||||||||||| |||||||||||||||||| ||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760


RESULT 3
SEPR_MOUSE
ID   SEPR_MOUSE     STANDARD;      PRT;   761 AA.
AC   P97321;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Seprase (EC 3.4.21.-) (Fibroblast activation protein alpha) (Integral
DE   membrane serine protease).
GN   Name=Fap;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE (ISOFORMS 1; 2 AND 3).
RC   STRAIN=BALB/c; TISSUE=Embryo;
RX   MEDLINE=97284459; PubMed=9139873;
RA   Niedermeyer J., Scanlan M.J., Garin-Chesa P., Daiber C., Fiebig H.H.,
RA   Old L.J., Rettig W.J., Schnapp A.;
RT   "Mouse fibroblast activation protein: molecular cloning, alternative
RT   splicing and expression in the reactive stroma of epithelial
RT   cancers.";
```

```
RL   Int. J. Cancer 71:383-389(1997).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC   TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May have a role in tissue remodeling during development
CC       and wound healing, and contribute to invasiveness in malignant
CC       cancers.
CC   -!- CATALYTIC ACTIVITY: Degrades gelatin and heat-denatured type I and
CC       type IV collagen, but not native type I or type IV collagen. Does
CC       not cleave laminin, fibronectin, fibrin or casein.
CC   -!- SUBUNIT: Homodimer, or heterodimer with DPP4. The monomer is
CC       inactive (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=1;
CC         IsoId=P97321-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P97321-2; Sequence=VSP_005368;
CC       Name=3;
CC         IsoId=P97321-3; Sequence=VSP_005369;
CC   -!- TISSUE SPECIFICITY: Detected in fibroblasts, in placenta, uterus,
CC       embryos from day 7-19 and in new-born mice (P1).
CC   -!- SIMILARITY: Belongs to the peptidase S9B family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; Y10007; CAA71116.1; -; mRNA.
DR   EMBL; BC019190; AAH19190.1; -; mRNA.
DR   HSSP; P27487; 1PFQ.
DR   MEROPS; S09.007; -.
DR   Ensembl; ENSMUSG00000000392; Mus musculus.
DR   MGI; MGI:109608; Fap.
DR   GO; GO:0016021; C:integral to membrane; TAS.
DR   InterPro; IPR002471; Pept_S9_AS.
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR002469; Peptidase_S9B.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF00930; DPPIV_N; 1.
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW   Alternative splicing; Glycoprotein; Hydrolase; Protease;
KW   Serine protease; Signal-anchor; Transmembrane.
FT   TOPO_DOM      1      4       Cytoplasmic (Potential).
FT   TRANSMEM      5     25       Signal-anchor for type II membrane
FT                                protein (Potential).
FT   TOPO_DOM     26    761       Extracellular (Potential).
FT   ACT_SITE    624    624       Charge relay system (By similarity).
FT   ACT_SITE    702    702       Charge relay system (By similarity).
FT   ACT_SITE    734    734       Charge relay system (By similarity).
FT   CARBOHYD     49     49       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     92     92       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     99     99       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    227    227       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    314    314       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    679    679       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC     31     63       Missing (in isoform 3).
FT                                /FTId=VSP_005369.
FT   VARSPLIC     31     35       Missing (in isoform 2).
FT                                /FTId=VSP_005368.
FT   CONFLICT    737    737       S -> L (in Ref. 2).
SQ   SEQUENCE   761 AA;  87945 MW;  9174C3AEDA213B25 CRC64;

  Query Match             90.7%;  Score 3705.5;  DB 1;  Length 761;
  Best Local Similarity   89.5%;  Pred. No. 6.3e-254;
  Matches  681;  Conservative   37;  Mismatches   42;  Indels    1;  Gaps    1;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||:| |||| | | |||:|:||||||||:  | || |||||||||||||||:|||
Db          1 MKTWLKTVFGVTTLAALALVVICIVLRPSRVYKPEGNTKRALTLKDILNGTFSYKTYFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||| ||||||| |:||| ||||| :|| |||| |||||||::||||||||||||||||||
Db         61 WISEQEYLHQSEDDNIVFYNIETRESYIILSNSTMKSVNATDYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              |||||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLQNGEFVRGYELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              ||||: ||||:|||||||||||||||||||||||:|||||| ||||:|||:|||||||
Db        181 FQITYTGRENRIFNGIPDWVYEEEMLATKYALWWSPDGKFLAYVEFNDSDIPIIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              | |||||||||||||||||||:||:||||| :||| ||||| |||||||||||||||:
Db        241 DGQYPRTINIPYPKAGAKNPVVRVFIVDTTYPHHVGPMEVPVPEMIASSDYYFSWLTWVS 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              ||||||||||||||||||||||||  |:||| |||:|||||||||||||| || |
Db        301 SERVCLQWLKRVQNVSVLSICDFREDWHAWECPKNQEHVEESRTGWAGGFFVSTPAFSQD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              | |||||||||||||||||||||||||||||||||||||| |||||||||||||||| |||
Db        361 ATSYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAIYIFRVTQDSLFYSSNEFEGYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||: ||||||||||||||||||||||||  ||||||||||:|||||||||
Db        421 RRNIYRISIGNSPPSKKCVTCHLRKERCQYYTASFSYKAKYYALVCYGPGLPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              |||||::|||||||||:|:|||||| |||||:   :| ||||||||||||||||||||||
Db        481 TDQEIQVLEENKELENSLRNIQLPKVEIKKLKDGGLTFWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              |||||||||:||||||||:|||||||:|||||||||||||||| |:|||||||||||||
Db        541 YGGPCSQSVKSVFAVNWITYLASKEGIVIALVDGRGTAFQGDKFLHAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              :|||||||||||||:|||||||||||||||||||||||||||||||||||||||||||:|
Db        601 LTAVRKFIEMGFIDEERIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASIY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 SERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||: || | ||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGISSGRSQNHLYTHMTHFLKQCFSLSD 761


RESULT 4
```

```
Q8R492_RAT
ID   Q8R492_RAT  PRELIMINARY;      PRT;   761 AA.
AC   Q8R492;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Fibroblast activation protein alpha subunit.
GN   Name=Fap;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Zuckschwerdt K., Park J.E.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF493782; AAM11677.1; -; mRNA.
DR   HSSP; P27487; 1PFQ.
DR   MEROPS; S09.007; -.
DR   Ensembl; ENSRNOG00000005679; Rattus norvegicus.
DR   RGD; 621253;. Fap.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004177; F:aminopeptidase activity; IEA.
DR   GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR002471; Pept_S9_AS.
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR002469; Peptidase_S9B.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF00930; DPPIV_N; 1.
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW   Hydrolase; Transmembrane.
SQ   SEQUENCE   761 AA;  87843 MW;  EB214BBE4BB14F2C CRC64;

  Query Match             90.2%;  Score 3686.5;  DB 2;  Length 761;
  Best Local Similarity   88.8%;  Pred. No. 1.4e-252;
  Matches  676;  Conservative   40;  Mismatches   44;  Indels    1;  Gaps    1;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||:| |||| |   |||:|:|||||||:: | || |:|||||||||||||||:|||
Db          1 MKTWLKTVFGVTTLVALALVVICIVLRPSRVYSPEGNTGRSLTLKDILNGTFSYKTYFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||| ||||||| |:||| ||||| :|| |||| |||||||::|||||||||:||||||||
Db         61 WISEQEYLHQSEDDNIVFYNIETRESYIILSNSTMKSVNATDYGLSPDRQFIYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              |||||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLQNGEFVRGYELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              ||||: ||||:|||||||||||||||||||||||:||:||| ||||:|||:|||||||
Db        181 FQITYTGRENRIFNGIPDWVYEEEMLATKYALWWSPDGKYLAYVEFNDSDIPIIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              | ||||||||||||||||||:||:||:|| || :||| ||||| ||||||||:||||||
Db        241 DGQYPRTINIPYPKAGAKNPIVRVFIVDTIYPHHVGPIEVPVPEMIASSDYYFTWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              :|||||||||||||||||||||||| ||||| |||||||||||||||||| || |
Db        301 NERVCLQWLKRVQNVSVLSICDFREDWHAWDCPKNQEHIEESRTGWAGGFFVSTPAFSQD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              | |||||||||||||||||||||||||||||||||||| |||||||||||||||| |||
Db        361 AASYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAIYIFRVTQDSLFYSSNEFEGYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||: |||||||||||||||||||||||  ||||||:|||||:|||||||||
Db        421 RRNIYRISIGNSPPSKKCVTCHLRKERCQYYTASFSYKAKYYALICYGPGLPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              |||||::|||||||||||:|||||  ||||||   :| ||||||||||||||||||||||
Db        481 TDQEIQVLEENKELENALRNIQLPAVEIKKLEDGGMTFWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              |||||||||:|||:||||:|||||||:|:|||||||||||||| |:|||||||||||||
Db        541 YGGPCSQSVKSVFSVNWITYLASKEGIVVALVDGRGTAFQGDKFLHAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              :|||||||||||||| ||||||||||||||||||||||||||||||||||||||||||:|
Db        601 LTAVRKFIEMGFIDEGRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASIY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||: || | ||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGISSGRSQNHLYTHMTHFLKQCFSLSD 761


RESULT 5
Q91651_XENLA
ID   Q91651_XENLA PRELIMINARY;      PRT;   755 AA.
AC   Q91651;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Fibroblast activation factor alpha.
GN   Name=gene 13;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=96312514; PubMed=8700860; DOI=10.1073/pnas.93.5.1924;
RA   Brown D.D., Wang Z., Furlow J.D., Kanamori A., Schwartzman R.A.,
RA   Remo B.F., Pinder A.;
RT   "The thyroid hormone-induced tail resorption program during Xenopus
RT   laevis metamorphosis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:1924-1929(1996).
DR   EMBL; U41856; AAC59872.1; -; mRNA.
DR   HSSP; P27487; 1PFQ.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004177; F:aminopeptidase activity; IEA.
DR   GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR002471; Pept_S9_AS.
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR002469; Peptidase_S9B.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF00930; DPPIV_N; 1.
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW   Hydrolase; Transmembrane.
SQ   SEQUENCE   755 AA;  86919 MW;  E07F41FEB6CB3DF9 CRC64;

  Query Match             60.8%;  Score 2485.5;  DB 2;  Length 755;
  Best Local Similarity   58.4%;  Pred. No. 2.3e-167;
  Matches  443;  Conservative  128;  Mismatches  184;  Indels    3;  Gaps    1;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||| |||:|||  : |: ||:: :||  |   | |   ::| :  |  :  |||: :
Db          1 MKTPVKILFGVLAATVVTLLIVTVVLVTSGFSNGE---LKAFSFDDAFSHNFSYQLYHHL 57

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              |||  ||: ::::||:::||:|||::  || | :  ||||:|| :| ||:|: |||:| |
Db         58 WISENEYIRKTSENNLIVYNVETGETSVILGNASFNSVNATNYIISSDRKFLLLESNYKK 117
```

```
Qy        121 LWRYSYTATYYIYDLSNGEPVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
Db        118 QWRYSYTASYHIYDIENGRPVTENELPPNIQYIYWAPKGHKLVYVLDNNIYLVYEPGQTA 177

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
Db        178 VALTNDGKENEIYNGIPDWVYEELMLGGKYATWWSPNGKFIAYVQFNDTEVPTIEYSFYG 237

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
Db        238 EDQYPRTMAIPYPKAGMKNPTVNLFIVKVDDPDSISTELIQVPSLLASSDYYFSWVTWVL 297

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
Db        298 DERLSVQWLPRSQNISVISICDFKKSSNTWVCPKNREHFDISETGWIGVFFPPLPVFTSD 357

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
Db        358 NLSYYQIISDRTGYKHIHHISDSKEKAIPITHGNWEAFYIYLVTDEAIYYTSTEYEGAPG 417

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
Db        418 RRHIYRIGLQGHLREKQCFTCNIRKERCQYYTARFSHKAKYYTLYCYGPGLPIVTLHEGS 477

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
Db        478 NNREIRILEDNADLEATLHNFEMPSEKIGKIVQDDKTFWYKMTLPPHFNARKKYPLLIYV 537

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
Db        538 YGGPSSQEVKPSFSVGYSTYLASTEGIIVASLDGRGTAYQGVIVMHAINNRLGTVAVEDQ 597

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
Db        598 IFAARKFIAMGFIDEKRIAIWGWSYGGYVTSMVLGPRPRPVKCGMTLPPVSNWEYYASIY 657

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
Db        658 TERYMGLPTKSDNLENYKNSTVMTRAEQFRKVDYLLVHGTADDNVHFQQAAQISKALVEA 717

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSL 758
Db        718 QVDFQAMWYTDKDHSISGSAKKHLFTHMTHFLKNWFNL 755


RESULT 6
Q75S82_AGKHB
ID   Q75S82_AGKHB PRELIMINARY;      PRT;   751 AA.
AC   Q75S82;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Dipeptidylpeptidase 4b (EC 3.4.14.5).
GN   Name=DPP4b;
OS   Agkistrodon halys brevicaudus (Korean slamosa snake) (Gloydius halys
OS   brevicaudus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Gloydius.
OX   NCBI_TaxID=259325;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Venom gland;
RA   Ogawa Y., Yanoshita R., Murayama N., Higuchi S., Samejima Y.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB158225; BAD06333.1; -; mRNA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004177; F:aminopeptidase activity; IEA.
DR   GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR002471; Pept_S9_AS.
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR002469; Peptidase_S9B.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF00930; DPPIV_N; 1.
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW   Hydrolase; Transmembrane.
SQ   SEQUENCE   751 AA;  86155 MW;  DAF219B4FEE7629A CRC64;

  Query Match            56.9%;  Score 2324;  DB 2;  Length 751;
  Best Local Similarity  55.1%;  Pred. No. 6.5e-156;
  Matches  421;  Conservative  140;  Mismatches  183;  Indels   20;  Gaps    9;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
Db          1 MKTVVKCLLGLLALGVIITAIVVPVVL----LTRDDSDIRRKFSLEDYLSDEFQYKSYNL 56

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYS 119
Db         57 RWMSGHEYVYTN-QNNVLLYNIDDERESIVLSNDTLDSFNSSQAILSPDRKFALLQYSYE 115

Qy        120 KLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDP 179
Db        116 KVWRHSYTASYHIYDLNNRTKITENPLPTNIQYISWSPVGHKLAYVYRNNVYVKATPNAS 175

Qy        180 PFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYY 239
Db        176 PVQITENGAENKILNGLADWVYEEEMFGTHSALWWSPNGRFLAFAEINDTEVPVMEYSFY 235

Qy        240 GDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYYFS 294
Db        236 SEDTLQYPKTIKIPYPKAGAINPTIRLFVLDIS----LSPKNISEIVAPSSIISGDHYLS 291

Qy        295 WLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVST 354
Db        292 VVTWVTDERICVQWLRRIQNFSVLTICDYS---GAWHCPKEREHLEESKTGWVGRFQPSE 348

Qy        355 PVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNE 414
Db        349 PYFTSDKISYYRIISDSEGYKHIHY-TDSAGKVKPITSGKWEVISISAVTNNSLYFISNE 407

Qy        415 FEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPIS 474
Db        408 FEGRPGGRHLYKVDLKN-DLKKICITCNSKEEACQYFSVSFSTDSRYYKLNCYGPDLPYF 466

Qy        475 TLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKY 534
Db        467 TLQNSITDKAIKTLEDNNNLKNVLKEIQMPCKRLSNITLHGQTYWYQMILPPNFDESKKY 526

Qy        535 PLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGV 594
Db        527 PLLIDVYAGPCSQKADAAFRINWSTYLASSEGIIVASFDGRGSGFQGDKILHAIYRRLGT 586

Qy        595 YEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWE 654
Db        587 YEVEDQISAAKLFSEMSFVDKDRIAIWGWSYGGYVTSMVLGAGSDVFKCGIAVAPVSRWQ 646

Qy        655 YYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIA 714
Db        647 YYDSIYTERYMGLPEKNDNLNFYENSTVMARAKNFRTVDYLLIHGTADDNVHFQQAAQIS 706

Qy        715 KALVNAQVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSL 758
Db        707 KALVDAEVDFQAMWYTDKDHGIGGHAHSHIYQHMSHFMKQCFKL 750


RESULT 7
Q75S83_AGKHB
ID   Q75S83_AGKHB PRELIMINARY;      PRT;   751 AA.
AC   Q75S83;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
```

```
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Dipeptidylpeptidase 4a (EC 3.4.14.5).
GN   Name=DPP4a;
OS   Agkistrodon halys brevicaudus (Korean slamosa snake) (Gloydius halys
OS   brevicaudus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Gloydius.
OX   NCBI_TaxID=259325;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Venom gland;
RA   Ogawa Y., Yanoshita R., Murayama N., Higuchi S., Samejima Y.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB158224; BAD06332.1; -; mRNA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004177; F:aminopeptidase activity; IEA.
DR   GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR002471; Pept_S9_AS.
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR002469; Peptidase_S9B.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF00930; DPPIV_N; 1.
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW   Hydrolase; Transmembrane.
SQ   SEQUENCE   751 AA;  86127 MW;  6EA3F2D829BE97D4 CRC64;

  Query Match             56.9%;  Score 2324;  DB 2;  Length 751;
  Best Local Similarity   55.1%;  Pred. No. 6.5e-156;
  Matches  421;  Conservative  140;  Mismatches  183;  Indels   20;  Gaps    9;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              ||| || : | :|   ::  :|: :||    :  : :  |  :|:| |:  | ||::
Db          1 MKTVVKCLLGLLALGVIITAIVVPVVL----LTRDDSDIRRKFSLEDYLSDEFQYKSYNL 56

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYS 119
               |:|| ||:: :  ||::||||:  :   :||| |: | |:|   |||||:|  |:  |
Db         57 RWMSGHEYVYTN-QNNVLLYNIDDERESIVLSNDTLDSFNSSQAILSPDRKFALLQYSYE 115

Qy        120 KLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDP 179
              |:||:||||:|:||||:|   :  | ||  |||: ||||| ||||||:||:|:|  |
Db        116 KVWRHSYTASYHIYDLNNRTKITENPLPTNIQYISWSPVGHKLAYVYRNNVYVKATPNAS 175

Qy        180 PFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYY 239
              | ||| || |||| ||: |||||||  |  ||||||||:|||:|| |||::||: ||:|
Db        176 PVQITENGAENKILNGLADWVYEEEMFGTHSALWWSPNGRFLAFAEINDTEVPVMEYSFY 235

Qy        240 GDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYYFS 294
               ::  |||:|| |||||||| || :|:|::| :    : |:   |:  |: | | |:| |
Db        236 SEDTLQYPKTIKIPYPKAGAINPTIRLFVLDIS----LSPKNISEIVAPSSIISGDHYLS 291

Qy        295 WLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVST 354
              :|||||||:|:|||:|:|| |||:|||:      | ||| :||:|||:||| | |  |
Db        292 AVTWVTDERICVQWLRRIQNFSVLTICDYS---GAWHCPKEREHLEESKTGWVGRFQPSE 348

Qy        355 PVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNE 414
              | |: | ||||:| || :|||||||  |:      |||||| |:|  || :||:: |||
Db        349 PYFTSDKISYYRIISDSEGYKHIHY-TDSAGKVKPITSGKWEVISISAVTNNSLYFISNE 407

Qy        415 FEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPIS 474
              ||  || |::|:: : :    | |:||: ::| |||:: |||  ::|| | |||| :|
Db        408 FEGRPGGRHLYKVDLKN-DLKKICITCNSKEEACQYFSVSFSTDSRYYKLNCYGPDLPYF 466

Qy        475 TLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKY 534
              || :  ||: || ||:|  |:| || ||:| : :  :  :    | ||:|||| || ||||
Db        467 TLQNSITDKAIKTLEDNNNLKNVLKEIQMPCKRLSNITLHGQTYWYQMILPPNFDESKKY 526

Qy        535 PLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGV 594
              |||| || |||||   : | :|| :|||| ||:::|  ||||: |||||:|:|:||:||
Db        527 PLLIDVYAGPCSQKADAAFRINWSTYLASSEGIIVASFDGRGSGFQGDKILHAIYRRLGT 586

Qy        595 YEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWE 654
              |||||||:| : | || |:|: |||||||||||||:|: | :|: :||||||||||| |:
Db        587 YEVEDQISAAKLFSEMSFVDKDRIAIWGWSYGGYVTSMVLGAGSDVFKCGIAVAPVSRWQ 646

Qy        655 YYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIA 714
              || |:||||:|||| |:|||  |:|||||||: || |||||||||||||||| :|||:
Db        647 YYDSIYTERYMGLPEKNDNLNFYENSTVMARAKNFRTVDYLLIHGTADDNVHFQQAAQIS 706

Qy        715 KALVNAQVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSL 758
              ||||:|:||||||||:|::||: | : :|:| ||:||:|||| |
Db        707 KALVDAEVDFQAMWYTDKDHGIGGHAHSHIYQHMSHFMKQCFKL 750


RESULT 8
Q5ZI81_CHICK
ID   Q5ZI81_CHICK PRELIMINARY;      PRT;   759 AA.
AC   Q5ZI81;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=RCJMB04_29g21;
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CB; TISSUE=Bursa;
RA   Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA   Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,
RA   Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT   "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT   genefunction analysis.";
RL   Genome Biol. 6:R6-R6(2005).
DR   EMBL; AJ720903; CAG32562.1; -; mRNA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004177; F:aminopeptidase activity; IEA.
DR   GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR002471; Pept_S9_AS.
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR002469; Peptidase_S9B.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF00930; DPPIV_N; 1.
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW   Hydrolase; Hypothetical protein; Transmembrane.
SQ   SEQUENCE   759 AA;  86343 MW;  B9A0C38CCB8644AC CRC64;

  Query Match             55.7%;  Score 2276;  DB 2;  Length 759;
  Best Local Similarity   54.8%;  Pred. No. 1.7e-152;
  Matches  417;  Conservative  130;  Mismatches  208;  Indels    6;  Gaps    4;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||| :| : |:   ||:  ::   :   :      | ::    ||:: ||  : |||
Db          1 MKTLLKWLLGLVGVAVVITVIAVPLALLTGESIPESDSRSTYTLENYLNNDYVYKTHNLQ 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              |||| :|||:::: ||: :: ||| |  :| | |:    |:   ||||::|  |:  | |
Db         61 WISGNQYLHETSNGNILRFDAETGTSSVVLLNTTISIHEATTAILSPDQRFALLQYKYEK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              |||:||||:|:||| :    :    ||   ||: ||||| |||||: ||||:|  |   |
Db        121 LWRHSYTASYHIYDFNTSSILDDALLPNDTQYISWSPVGHKLAYVWNNNIYIKASPTAAP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              ||| || |||||||||||||||  :  |||||||| |:||| ||||::||| ||:|
```

```
Db        181 VQITSNGEENKIFNGIPDWVYEEEMFGSHSALWWSPNGNFVAYAAFNDTEVPVIEYSFYS 240

Qy        241 DE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTW 298
              ::  |||:|| |||||||||||| |: ||:|          |:  || | | |:| | :||
Db        241 EDTLQYPKTIRIPYPKAGAKNPTVKFFIVDIQMLPDFNSTEISPPAEIKSGDHYLSVVTW 300

Qy        299 VTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESR-TGWAGGFFVSTPVF 357
              ||||:|||||:|:|| |||:||||      | | :     :: :  ||         |:|
Db        301 VTDERICLQWLRRIQNYSVLTICDFESATGNWTCHRKNNFWKKVQLAGWQ--ISAICPLF 358

Qy        358 SYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEE 417
              : |  :|||:||: :|||||||  | |  : || |||| |:|  ||:  |:| ||:  |
Db        359 APDNTTYYKVFSNTEGYKHIHYINGT-EAPVPITEGKWEVISIAAVTKYFLYYISNQNGE 417

Qy        418 YPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLH 477
               || ||:|:: : | | | :||:| | :||||||:|||| |:|| | | |||:|:|||| 
Db        418 MPGGRNLYKMLLESSPKSTQCVSCDLNQERCQYYSASFSKDAQYYQLNCLGPGLPMSTLH 477

Qy        478 DGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLL 537
                 || :: || | ||||:||:||:| :::  : |    |||:|||||  | |||||||
Db        478 RSSDDQVLRYLENNTELENSLKDIQMPSKKLGSITVGGYNLWYQMILPPHLDSSKKYPLL 537

Qy        538 IQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEV 597
              ::|| ||||| |  || :|| :|||| | :::|  ||||: :|||:::::|: |:|| |||
Db        538 LEVYAGPCSQKVDHVFRINWATYLASTEQIIVASFDGRGSGYQGDEIMHAINRRLGTYEV 597

Qy        598 EDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYA 657
              ||||:| | | || |:|: |||||||||||||:|: | ||:|:|||||||||||| |:||
Db        598 EDQISAARTFSEMSFVDKDRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSRWQYYD 657

Qy        658 SVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKAL 717
              |:|||:|||||: ||| :| :|||||||| |: |:|||||||||||||| :|||:|||
Db        658 SIYTERYMGLPTESDNLRNYNSSTVMARAEKFKEVEYLLIHGTADDNVHFQQAAQISKAL 717

Qy        718 VNAQVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSL 758
              |:|:|||||||:|::||:|| :  |:||||:||:|||||
Db        718 VDAEVDFQAMWYTDKDHGISGQAHKHIYTHMSHFIKQCFSL 758
```

RESULT 9
DPP4_PIG
ID DPP4_PIG STANDARD; PRT; 766 AA.
AC P22411; Q866G2;
DT 01-AUG-1991 (Rel. 19, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
DE IV) (T-cell activation antigen CD26) [Contains: Dipeptidyl peptidase 4
DE membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl
DE peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)].
GN Name=DPP4; Synonyms=CD26;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND FUNCTION.
RC TISSUE=Kidney;
RX PubMed=14719797; DOI=10.1515/BC.2003.172;
RA Baer J., Weber A., Hoffmann T., Stork J., Wermann M., Wagner L.,
RA Aust S., Gerhartz B., Demuth H.-U.;
RT "Characterisation of human dipeptidyl peptidase IV expressed in Pichia
RT pastoris. A structural and mechanistic comparison between the
RT recombinant human and the purified porcine enzyme.";
RL Biol. Chem. 384:1553-1563(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 2-67.
RC TISSUE=Kidney;
RX MEDLINE=94093209; PubMed=7903569;
RA Thomsen P.D., Qvist H., Marklund L., Andersson L., Sjostrom H.,
RA Noren O.;
RT "Assignment of the dipeptidylpeptidase IV (DPP4) gene to pig
RT chromosome 15q21.";
RL Mamm. Genome 4:604-607(1993).
RN [3]
RP PROTEIN SEQUENCE OF 38-71.
RC TISSUE=Kidney;
RX MEDLINE=91273813; PubMed=1675855;
RA Seidl R., Mann K., Schaeffer W.;
RT "N-terminal amino-acid sequence of pig kidney dipeptidyl peptidase IV
RT solubilized by autolysis.";
RL Biol. Chem. Hoppe-Seyler 372:213-214(1991).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 39-766, N-GLYCOSYLATION
RP SITES, AND HOMODIMERIZATION.
RX PubMed=12690074; DOI=10.1073/pnas.0230620100;
RA Engel M., Hoffmann T., Wagner L., Wermann M., Heiser U.,
RA Kiefersauer R., Huber R., Bode W., Demuth H.-U., Brandstetter H.;
RT "The crystal structure of dipeptidyl peptidase IV (CD26) reveals its
RT functional regulation and enzymatic mechanism.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5063-5068(2003).
CC -!- FUNCTION: Removes N-terminal dipeptides sequentially from
CC polypeptides having unsubstituted N-termini provided that the
CC penultimate residue is proline.
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
CC soluble form.
CC -!- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC by proteolytic processing.
CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC ---
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC ---
DR EMBL; AY198323; AAO43404.1; -; mRNA.
DR EMBL; X73276; CAA51717.1; -; mRNA.
DR PIR; I47134; I47134.
DR PDB; 1ORV; X-ray; A/B/C/D=39-766.
DR PDB; 1ORW; X-ray; A/B/C/D=39-766.
DR SMR; P22411; 39-766.
DR MEROPS; S09.003; -.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW 3D-structure; Aminopeptidase; Direct protein sequencing; Glycoprotein;
KW Hydrolase; Protease; Serine protease; Signal-anchor; Transmembrane.

FT	CHAIN	1	766	Dipeptidyl peptidase 4 membrane form.
FT	CHAIN	38	766	Dipeptidyl peptidase 4 soluble form.
FT	TOPO_DOM	1	6	Cytoplasmic (Potential).
FT	TRANSMEM	7	27	Signal-anchor for type II membrane protein (Potential).
FT	TOPO_DOM	28	766	Extracellular (Potential).
FT	ACT_SITE	630	630	Charge relay system (By similarity).
FT	ACT_SITE	708	708	Charge relay system (By similarity).
FT	ACT_SITE	740	740	Charge relay system (By similarity).
FT	CARBOHYD	85	85	N-linked (GlcNAc. . .).
FT	CARBOHYD	92	92	N-linked (GlcNAc. . .).
FT	CARBOHYD	150	150	N-linked (GlcNAc. . .) (By similarity).
FT	CARBOHYD	179	179	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	219	219	N-linked (GlcNAc. . .) (By similarity).
FT	CARBOHYD	229	229	N-linked (GlcNAc. . .).
FT	CARBOHYD	279	279	N-linked (GlcNAc. . .).
FT	CARBOHYD	321	321	N-linked (GlcNAc. . .).
FT	CARBOHYD	393	393	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	685	685	N-linked (GlcNAc. . .).

```
FT   DISULFID    385    394
FT   DISULFID    444    447
FT   DISULFID    454    472
FT   DISULFID    649    762
FT   CONFLICT     32     32       Missing (in Ref. 2).
SQ   SEQUENCE   766 AA;  88242 MW;  8800D520BAEA856D CRC64;

  Query Match             54.1%;  Score 2208.5;  DB 1;  Length 766;
  Best Local Similarity   51.8%;  Pred. No. 1e-147;
  Matches  397;  Conservative  153;  Mismatches  206;  Indels   11;  Gaps    7;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  |:: | :  :|:: :: : :||      ::  :: |  || | |  ||  | :
Db          1 MKTPWKVLLGLLGIAALVTVITVPVVLLNKGTDDAAADSRRTYTLTDYLKSTFRVKFYTL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSV--NASNYGLSPDRQFVYLESD 117
               |||  |||::  :|||:|:| | | |   | | |   :  : ::| :||||||:  | :
Db         61 QWISDHEYLYKQ-ENNILLFNAEYGNSSIFLENSTFDELGYSTNDYSVSPDRQFILFEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWITWSPVGHKLAYVWNNDIYVKNEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
                :||: |:|| |:||: |||||||: :   |||||||| |||||:||||::|:|  ||
Db        180 LSSQRITWTGKENVIYNGVTDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDT---TYPAYVGPQEVPVPAMIASSDYY 292
              :| ||  |||:|: |||||||||:|| |: |::||   :  | |   ::  || :   |:|
Db        240 FYSDESLQYPKTVRIPYPKAGAENPTVKFFVVDTRTLSPNASVTSYQIVPPASVLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
                :||||:||: |||::| || |:: |||: |    |      ::||| | ||| | |
Db        300 LCGVTWVTEERISLQWIRRAQNYSIIDICDYDESTGRWISSVARQHIEISTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              : | |: |  |:||| |:::||||| : :    |   || | || | |  :| | |:| |
Db        360 AEPHFTSDGNSFYKIISNEEGYKHICHFQTDKSNCTFITKGAWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              || :  || ||:||| :  |     |::| |  |||||:|||| |||| | |:|||:|
Db        420 NEHKGMPGGRNLYRIQLNDY-TKVTCLSCELNPERCQYYSASFSNKAKYYQLRCFGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||   :|:|:::||:|  |:  |:::|:| :::  : :     ||:||||| ||:||
Db        479 LYTLHSSSSDKELRVLEDNSALDKMLQDVQMPSKKLDVINLHGTKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              ||||||:|| ||||| | :|| ::| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLIEVYAGPCSQKVDTVFRLSWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||| | |:| :|||:|:|||||||||||||:|: | :|:|:|||||||||||
Db        599 GTFEVEDQIEATRQFSKMGFVDDKRIAIWGWSYGGYVTSMVLGAGSGVFKCGIAVAPVSK 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| ||||||:||||| :|||::|:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDYYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              ::||||:| |||| |||:|::||: | ::  |:||||:||||||||
Db        719 LSKALVDAGVDFQTMWYTDEDHGIASNMAHQHIYTHMSHFLKQCFSL 765


RESULT 10
DPP4_RAT
ID   DPP4_RAT       STANDARD;      PRT;   767 AA.
AC   P14740;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-FEB-2005 (Rel. 46, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
DE   IV) (T-cell activation antigen CD26) (GP110 glycoprotein) (Bile
DE   canaliculus domain-specific membrane glycoprotein) [Contains:
DE   Dipeptidyl peptidase 4 membrane form (Dipeptidyl peptidase IV membrane
DE   form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV
DE   soluble form); Dipeptidyl peptidase 4 60 kDa soluble form (Dipeptidyl
DE   peptidase IV 60 kDa soluble form)].
GN   Name=Dpp4; Synonyms=Cd26;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
RX   MEDLINE=89123496; PubMed=2563382;
RA   Ogata S., Misumi Y., Ikehara Y.;
RT   "Primary structure of rat liver dipeptidyl peptidase IV deduced from
RT   its cDNA and identification of the NH2-terminal signal sequence as the
RT   membrane-anchoring domain.";
RL   J. Biol. Chem. 264:3596-3601(1989).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=88068516; PubMed=3479775;
RA   Hong W., Doyle D.;
RT   "cDNA cloning for a bile canaliculus domain-specific membrane
RT   glycoprotein of rat hepatocytes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:7962-7966(1987).
RN   [3]
RP   NUCLEOTIDE SEQUENCE OF 1-40.
RX   MEDLINE=89034185; PubMed=3182821;
RA   Hong W.J., Doyle D.;
RT   "Membrane orientation of rat gp110 as studied by in vitro
RT   translation.";
RL   J. Biol. Chem. 263:16892-16898(1988).
RN   [4]
RP   PROTEIN SEQUENCE OF 28-58, AND TISSUE SPECIFICITY.
RX   MEDLINE=90228896; PubMed=1970322;
RA   McCaughan G.W., Wickson J.E., Creswick P.F., Gorrell M.D.;
RT   "Identification of the bile canalicular cell surface molecule GP110 as
RT   the ectopeptidase dipeptidyl peptidase IV: an analysis by tissue
RT   distribution, purification and N-terminal amino acid sequence.";
RL   Hepatology 11:534-544(1990).
RN   [5]
RP   PROTEIN SEQUENCE OF 281-302, AND MUTAGENESIS OF GLY-629; TRP-630;
RP   SER-631; TRY-632 AND GLY-633.
RC   TISSUE=Kidney;
RX   MEDLINE=94128239; PubMed=7905271;
RA   Iwaki-Egawa S., Watanabe Y., Fujimoto Y.;
RT   "N-terminal amino acid sequence of the 60-kDa protein of rat kidney
RT   dipeptidyl peptidase IV.";
RL   Biol. Chem. Hoppe-Seyler 374:973-975(1993).
RN   [6]
RP   PROTEIN SEQUENCE OF 624-648.
RX   MEDLINE=92190188; PubMed=1347701;
RA   Ogata S., Misumi Y., Tsuji E., Takami N., Oda K., Ikehara Y.;
RT   "Identification of the active site residues in dipeptidyl peptidase IV
RT   by affinity labeling and site-directed mutagenesis.";
RL   Biochemistry 31:2582-2587(1992).
RN   [7]
RP   SIGNAL-ANCHOR.
RX   MEDLINE=90338089; PubMed=1974258; DOI=10.1083/jcb.111.2.323;
RA   Hong W., Doyle D.;
RT   "Molecular dissection of the NH2-terminal signal/anchor sequence of
RT   rat dipeptidyl peptidase IV.";
RL   J. Cell Biol. 111:323-328(1990).
CC   -!- FUNCTION: Removes N-terminal dipeptides sequentially from
CC       polypeptides having unsubstituted N-termini provided that the
CC       penultimate residue is proline.
CC   -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC       Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC       Zaa is neither Pro nor hydroxyproline.
CC   -!- SUBUNIT: Homodimer.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
CC       soluble form.
```

```
CC   -!- TISSUE SPECIFICITY: Expressed in bile ducts and other epithelial
CC       brush borders (small intestine, kidney, colon, pancreatic duct);
CC       acinar structures in salivary glands; endothelial structures and T
CC       cell areas in thymus, spleen and lymph node.
CC   -!- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC       by proteolytic processing.
CC   -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; J04591; AAA41096.1; -; mRNA.
DR   EMBL; J02997; AAA41272.1; -; mRNA.
DR   PIR; A39914; A39914.
DR   HSSP; P27487; 1PFQ.
DR   SMR; P14740; 37-766.
DR   MEROPS; S09.003; -.
DR   Ensembl; ENSRNOG00000030763; Rattus norvegicus.
DR   RGD; 2515; Dpp4.
DR   GO; GO:0005624; C:membrane fraction; IDA.
DR   GO; GO:0005625; C:soluble fraction; IDA.
DR   GO; GO:0004274; F:dipeptidyl-peptidase IV activity; TAS.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
DR   InterPro; IPR002471; Pept_S9_AS.
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR002469; Peptidase_S9B.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF00930; DPPIV_N; 1.
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW   Aminopeptidase; Direct protein sequencing; Glycoprotein; Hydrolase;
KW   Protease; Serine protease; Signal-anchor; Transmembrane.
FT   CHAIN         1    767       Dipeptidyl peptidase 4 membrane form.
FT   CHAIN        37    767       Dipeptidyl peptidase 4 soluble form.
FT   CHAIN       281    767       Dipeptidyl peptidase 4 60 kDa soluble
FT                                form.
FT   TOPO_DOM      1      6       Cytoplasmic (Potential).
FT   TRANSMEM      7     28       Signal-anchor for type II membrane
FT                                protein (Potential).
FT   TOPO_DOM     29    767       Extracellular (Potential).
FT   ACT_SITE    631    631       Charge relay system (By similarity).
FT   ACT_SITE    709    709       Charge relay system (By similarity).
FT   ACT_SITE    741    741       Charge relay system (By similarity).
FT   CARBOHYD     83     83       N-linked (GlcNAc. . .) (By similarity).
FT   CARBOHYD     90     90       N-linked (GlcNAc. . .) (By similarity).
FT   CARBOHYD    148    148       N-linked (GlcNAc. . .) (By similarity).
FT   CARBOHYD    217    217       N-linked (GlcNAc. . .) (By similarity).
FT   CARBOHYD    227    227       N-linked (GlcNAc. . .) (By similarity).
FT   CARBOHYD    319    319       N-linked (GlcNAc. . .) (By similarity).
FT   CARBOHYD    521    521       N-linked (GlcNAc. . .) (By similarity).
FT   CARBOHYD    686    686       N-linked (GlcNAc. . .) (By similarity).
FT   DISULFID    383    395       By similarity.
FT   DISULFID    445    448       By similarity.
FT   DISULFID    455    473       By similarity.
FT   DISULFID    650    763       By similarity.
FT   MUTAGEN     629    629       G->A: Reduced activity.
FT   MUTAGEN     629    629       G->R: Reduced activity.
FT   MUTAGEN     630    630       W->E: No effect on activity.
FT   MUTAGEN     631    631       S->A: Reduced activity.
FT   MUTAGEN     632    632       Y->F: No effect on activity.
FT   MUTAGEN     632    632       Y->G: Reduced activity.
FT   MUTAGEN     632    632       Y->L: Reduced activity.
FT   MUTAGEN     633    633       G->A: Reduced activity.
FT   MUTAGEN     633    633       G->S: Reduced activity.
FT   CONFLICT     38     38       R -> A (in Ref. 1).
FT   CONFLICT     54     54       Missing (in Ref. 4).
FT   CONFLICT    183    183       I -> T (in Ref. 2).
FT   CONFLICT    332    332       T -> N (in Ref. 2).
FT   CONFLICT    352    352       C -> V (in Ref. 2).
FT   CONFLICT    394    394       V -> D (in Ref. 2).
FT   CONFLICT    562    562       L -> F (in Ref. 2).
FT   CONFLICT    624    624       R -> Q (in Ref. 2).
SQ   SEQUENCE   767 AA;  88089 MW;  ED947174F1F3E440 CRC64;

  Query Match             53.9%;  Score 2203;  DB 1;  Length 767;
  Best Local Similarity   52.4%;  Pred. No. 2.6e-147;
  Matches  405;  Conservative  148;  Mismatches  198;  Indels   22;  Gaps   11;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEE---NTMRALTLKDILNGTFSYKT 56
              |||  |:: | :  :|:: :: : :||      | :|   :: |  || | |  ||  |:
Db          1 MKTPWKVLLGLLGVAALVTIITVPVVLL-----NKDEAAADSRRTYTLADYLKNTFRVKS 55

Qy         57 FFPNWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMK--SVNASNYGLSPDRQFVYL 114
              :   |:|  |||::  :|||:|:| | | |   | | | :    : |:| :|||| || |
Db         56 YSLRWVSDSEYLYKQ-ENNILLFNAEHGNSSIFLENSTFEIFGDSISDYSVSPDRLFVLL 114

Qy        115 ESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQ 174
              | :| | ||:||||:| ||||:  : :   ::|   |:: ||  | |||||::|:||:|
Db        115 EYNYVKQWRHSYTASYSIYDLNKRQLITEEKIPNNTQWITWSQEGHKLAYVWKNDIYVKI 174

Qy        175 RPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVI 234
              |  | :||  |:|| ||||| |||||||:     |||||||| |||||:|||| :|:|
Db        175 EPHLPSHRITSTGKENVIFNGINDWVYEEEIFGAYSALWWSPNGTFLAYAQFNDTGVPLI 234

Qy        235 AYSYYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDT---TYPAYVGPQEVPVPAMIASS 289
              ||:| ||  |||:|: |||||||| || |: ||::|   :       | ::  || : :
Db        235 EYSFYSDESLQYPKTVWIPYPKAGAVNPTVKFFIVNTDSLSSTTTTIPMQITAPASVTTG 294

Qy        290 DYYFSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGG 349
              |:|   : ||:::|: |||||:|:|| ||::|||: :    |:|| |||||| | ||| |
Db        295 DHYLCDVAWVSEDRISLQWLRRIQNYSVMAICDYDKTTLVWNCPTTQEHIETSATGWCGR 354

Qy        350 FFVSTPVFSYDAISYYKIFSDKDGYKHI-HYIKDTVENAI--QITSGKWEAINIFRVTQD 406
              |  : | |: |  |:||| |||||||| : ||     :   || | || |:|  :| |
Db        355 FRPAEPHFTSDGSSFYKIVSDKDGYKHICQFQKDRKPEQVCTFITKGAWEVISIEALTSD 414

Qy        407 SLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVC 466
              |:| |||::| || ||:|:| :  :  :|||::| |  ||||||: | |  |||| | |
Db        415 YLYYISNEYKEMPGGRNLYKIQLTDH-TNKKCLSCDLNPERCQYYSVSLSKEAKYYQLGC 473

Qy        467 YGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPP 526
              |||:|: |||     :|:::||:|  |:  |:::|:| :::  : ::|   ||:|||||
Db        474 RGPGLPLYTLHRSTDQKELRVLEDNSALDKMLQDVQMPSKKLDFIVLNETRFWYQMILPP 533

Qy        527 QFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLY 586
              ||:|||||||| || |||||   : | :|| :|||| | :::|  ||||: :||||:::
Db        534 HFDKSKKYPLLIDVYAGPCSQKADAAFRLNWATYLASTENIIVASFDGRGSGYQGDKIMH 593

Qy        587 AVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIA 646
              |: ::||  |||||| | |:|::|||:| ||:||||||||||:|: | ||:|:||||||
Db        594 AINKRLGTLEVEDQIEAARQFLKMGFVDSKRVAIWGWSYGGYVTSMVLGSGSGVFKCGIA 653

Qy        647 VAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVH 706
              ||||| |||| |||||:|||| :|||:||:|||||:||| |: |:|||||||||||||
Db        654 VAPVSRWEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVH 713

Qy        707 FQNSAQIAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              || ||||:||||:| ||||||||:|::||: |  :  |:|:||:|||:|||||
Db        714 FQQSAQISKALVDAGVDFQAMWYTDEDHGIASSTAHQHIYSHMSHFLQQCFSL 766


RESULT 11
DPP4_HUMAN
ID   DPP4_HUMAN     STANDARD;      PRT;   766 AA.
AC   P27487;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
DE   IV) (T-cell activation antigen CD26) (TP103) (Adenosine deaminase
DE   complexing protein 2) (ADABP) [Contains: Dipeptidyl peptidase 4
DE   membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl
DE   peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)].
```

```
GN   Name=DPP4; Synonyms=ADCP2, CD26;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RX   MEDLINE=92329551; PubMed=1352704; DOI=10.1016/0167-4781(92)90036-Y;
RA   Misumi Y., Hayashi Y., Arakawa F., Ikehara Y.;
RT   "Molecular cloning and sequence analysis of human dipeptidyl peptidase
RT   IV, a serine proteinase on the cell surface.";
RL   Biochim. Biophys. Acta 1131:333-336(1992).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Placenta;
RX   MEDLINE=95012454; PubMed=7927537;
RA   Abbott C.A., Baker E., Sutherland G.R., McCaughan G.W.;
RT   "Genomic organization, exact localization, and tissue expression of
RT   the human CD26 (dipeptidyl peptidase IV) gene.";
RL   Immunogenetics 40:331-338(1994).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Peripheral blood;
RX   MEDLINE=92325476; PubMed=1352530;
RA   Tanaka T., Camerini D., Seed B., Torimoto Y., Dang N.H., Kameoka J.,
RA   Dahlberg H.N., Schlossman S.F., Morimoto C.;
RT   "Cloning and functional expression of the T cell activation antigen
RT   CD26.";
RL   J. Immunol. 149:481-486(1992).
RN   [4]
RP   ERRATUM.
RX   MEDLINE=93171637; PubMed=8094732;
RA   Tanaka T.;
RL   J. Immunol. 150:2090-2090(1993).
RN   [5]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Prostate, and Uterus;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   NUCLEOTIDE SEQUENCE OF 1-551.
RC   TISSUE=Colon;
RX   MEDLINE=92165847; PubMed=1347043;
RA   Darmoul D., Lacasa M., Baricault L., Marguet D., Sapin C., Trotot P.,
RA   Barbat A.;
RT   "Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like
RT   colon cancer cell lines HT-29 and Caco-2. Cloning of the complete
RT   human coding sequence and changes of dipeptidyl peptidase IV mRNA
RT   levels during cell differentiation.";
RL   J. Biol. Chem. 267:4824-4833(1992).
RN   [7]
RP   NUCLEOTIDE SEQUENCE OF 545-766.
RC   TISSUE=Colon;
RX   MEDLINE=91024044; PubMed=1977364;
RA   Darmoul D., Lacasa M., Chantret I., Swallow D., Trugnan G.;
RT   "Isolation of a cDNA probe for the human intestinal
RT   dipeptidylpeptidase IV and assignment of the gene locus DPP4 to
RT   chromosome 2.";
RL   Ann. Hum. Genet. 54:191-197(1990).
RN   [8]
RP   NUCLEOTIDE SEQUENCE OF 1-31.
RX   MEDLINE=96067599; PubMed=7487939;
RA   Boehm S.K., Gum J.R. Jr., Erickson R.H., Hicks J.W., Kim Y.S.;
RT   "Human dipeptidyl peptidase IV gene promoter: tissue-specific
RT   regulation from a TATA-less GC-rich sequence characteristic of a
RT   housekeeping gene promoter.";
RL   Biochem. J. 311:835-843(1995).
RN   [9]
RP   PROTEIN SEQUENCE OF 1-22, AND TISSUE SPECIFICITY.
RX   PubMed=1677636;
RA   Gorvel J.P., Ferrero A., Chambraud L., Rigal A., Bonicel J.,
RA   Maroux S.;
RT   "Expression of sucrase-isomaltase and dipeptidylpeptidase IV in human
RT   small intestine and colon.";
RL   Gastroenterology 101:618-625(1991).
RN   [10]
RP   PARTIAL PROTEIN SEQUENCE.
RC   TISSUE=Kidney;
RX   MEDLINE=93210468; PubMed=8096237; DOI=10.1084/jem.177.4.1135;
RA   Morrison M.E., Vijayasaradhi S., Engelstein D., Albino A.P.,
RA   Houghton A.N.;
RT   "A marker for neoplastic progression of human melanocytes is a cell
RT   surface ectopeptidase.";
RL   J. Exp. Med. 177:1135-1143(1993).
RN   [11]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 38-766 IN COMPLEX WITH
RP   INHIBITOR, AND HOMODIMERIZATION.
RX   PubMed=12832764; DOI=10.1107/S0907444903010059;
RA   Oefner C., D'Arcy A., Mac Sweeney A., Pierau S., Gardiner R.,
RA   Dale G.E.;
RT   "High-resolution structure of human apo dipeptidyl peptidase IV/CD26
RT   and its complex with 1-[([2-[(5-iodopyridin-2-yl)amino]-ethyl]amino)-
RT   acetyl]-2-cyano-(S)-pyrrolidine.";
RL   Acta Crystallogr. D 59:1206-1212(2003).
RN   [12]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-771, HOMODIMERIZATION, AND
RP   N-GLYCOSYLATION SITES.
RX   PubMed=12646248; DOI=10.1016/S0006-291X(03)00258-4;
RA   Hiramatsu H., Kyono K., Higashiyama Y., Fukushima C., Shima H.,
RA   Sugiyama S., Inaka K., Yamamoto A., Shimizu R.;
RT   "The structure and function of human dipeptidyl peptidase IV,
RT   possessing a unique eight-bladed beta-propeller fold.";
RL   Biochem. Biophys. Res. Commun. 302:849-854(2003).
RN   [13]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND
RP   N-GLYCOSYLATION SITES.
RX   PubMed=12483204; DOI=10.1038/nsb882;
RA   Rasmussen H.B., Branner S., Wiberg F.C., Wagtmann N.;
RT   "Crystal structure of human dipeptidyl peptidase IV/CD26 in complex
RT   with a substrate analog.";
RL   Nat. Struct. Biol. 10:19-25(2003).
RN   [14]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND
RP   N-GLYCOSYLATION SITES.
RX   PubMed=12906826; DOI=10.1016/S0969-2126(03)00160-6;
RA   Thoma R., Loeffler B., Stihle M., Huber W., Ruf A., Hennig M.;
RT   "Structural basis of proline-specific exopeptidase activity as
RT   observed in human dipeptidyl peptidase-IV.";
RL   Structure 11:947-959(2003).
CC   -!- FUNCTION: Removes N-terminal dipeptides sequentially from
CC       polypeptides having unsubstituted N-termini provided that the
CC       penultimate residue is proline. Plays a role in T cell activation.
CC   -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC       Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC       Zaa is neither Pro nor hydroxyproline.
```

```
CC   -!- SUBUNIT: Homodimer or heterodimer with Seprase (FAP).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
CC       soluble form.
CC   -!- TISSUE SPECIFICITY: Expressed in the poorly differentiated crypt
CC       cells of the small intestine as well as in the mature villous
CC       cells. Expressed at very low levels in the colon.
CC   -!- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC       by proteolytic processing.
CC   -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; U13735; AAB60646.1; -; Genomic_DNA.
DR   EMBL; U13710; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13711; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13712; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13713; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13714; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13715; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13716; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13717; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13718; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13719; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13720; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13721; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13722; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13723; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13724; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13725; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13726; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13727; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13728; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13729; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13730; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13731; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13732; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13733; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; U13734; AAB60646.1; JOINED; Genomic_DNA.
DR   EMBL; M74777; AAA51943.1; -; mRNA.
DR   EMBL; BC013329; AAH13329.2; -; mRNA.

  Query Match             53.6%;  Score 2190.5;  DB 1;  Length 766;
  Best Local Similarity   51.5%;  Pred. No. 2e-146;
  Matches  395;  Conservative  149;  Mismatches  212;  Indels   11;  Gaps    7;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  |:: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db          1 MKTPWKVLLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
               |||  |||::  :|||:::| | | |   | | |       |  :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
               | ::||: |:|: |:||| |||||||: :   |||||||| |||||:||||::|:| ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :|||||| || |: |:::|   : |      ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
                :|| | ||: ||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||::  || ||:|:| :  |      |::| |  |||||: ||| |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||    |: :::||:|  |:  |:|:|:| :::  : ::|   ||:|||| ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||| | |:| :|||:| |||||||||||||:|: | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| ||||||:||||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |:||||:  |||||||:|::||: |  :  |:||||:||:||||||
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765


RESULT 12
Q53TN1_HUMAN
ID   Q53TN1_HUMAN PRELIMINARY;      PRT;   766 AA.
AC   Q53TN1;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein DPP4.
GN   Name=DPP4;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Ozanich A., Stoneking T., Hawkins M., Sapetti L.;
RT   "The sequence of Homo sapiens BAC clone RP11-178A14.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Waterston R.H.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RA   Waterston R.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RA   Wilson R.K.;
RL   Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC008063; AAX93179.1; -; Genomic_DNA.
DR   SMR; Q53TN1; 39-766.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0003824; F:catalytic activity; IEA.
DR   GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR   GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR002471; Pept_S9_AS.
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR002469; Peptidase_S9B.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF00930; DPPIV_N; 1.
DR   Pfam; PF00326; Peptidase_S9; 1.
```

```
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW   Hydrolase; Hypothetical protein; Transmembrane.
SQ   SEQUENCE   766 AA;  88279 MW;  5FB4A2C6662D6117 CRC64;

  Query Match             53.6%;  Score 2190.5;  DB 2;  Length 766;
  Best Local Similarity   51.5%;  Pred. No. 2e-146;
  Matches  395;  Conservative  149;  Mismatches  212;  Indels   11;  Gaps    7;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              |||  |:: | :  :|:: :: : :||      ::  :: :  || | |  |:  | :
Db          1 MKTPWKVLLGLLGAAALVTIITVPVVLLNKGTDDATADSRKTYTLTDYLKNTYRLKLYSL 60

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESD 117
               |||  |||::  :|||:::| | | |   | | |      |  :| :||| ||: || :
Db         61 RWISDHEYLYKQ-ENNILVFNAEYGNSSVFLENSTFDEFGHSINDYSISPDGQFILLEYN 119

Qy        118 YSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPG 177
              | | ||:||||:| ||||:  : :    :|   |:: ||||| |||||: |:||:|  |
Db        120 YVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPN 179

Qy        178 DPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYS 237
              | ::||: |:|: |:||| |||||||: :   |||||||| |||||:||||::|:|  ||
Db        180 LPSYRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYS 239

Qy        238 YYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYY 292
              :| ||  |||:|: :||||||| || |: |:::|   : |       ::  || :   |:|
Db        240 FYSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHY 299

Qy        293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFV 352
               :|| | ||: ||||:|:|| ||: |||: |    |:|   ::||| | ||| | |
Db        300 LCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRP 359

Qy        353 STPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSS 412
              | | |: |  |:||| |:::||:|| | :   ::   || | || | |  :| | |:| |
Db        360 SEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYIS 419

Qy        413 NEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIP 472
              ||:  || ||:|:| :  |     |::| |  |||||: ||| |||| | | |||:|
Db        420 NEYKGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLP 478

Qy        473 ISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSK 532
              : |||     |: :::||:|  |:  |:|:|:| :::  : ::|   ||:|||| ||:||
Db        479 LYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSK 538

Qy        533 KYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
              |||||: || |||||   :|| :|| :|||| | :::|  ||||: :||||:::|: |:|
Db        539 KYPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRL 598

Qy        593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSS 652
              | :|||||| | |:| :|||:| ||||||||||||||:|: | ||:|:|||||||||||
Db        599 GTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSR 658

Qy        653 WEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQ 712
              |||| |||||:|||| :|||:||:|||||:||| |: |:|||||||||||||| |||
Db        659 WEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQ 718

Qy        713 IAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              |:||||:  |||||||:|::||: |  :  |:||||:||:||||||
Db        719 ISKALVDVGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 765


RESULT 13
P70092_XENLA
ID   P70092_XENLA PRELIMINARY;      PRT;   748 AA.
AC   P70092;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Dipeptidyl-peptidase IV (EC 3.4.14.5).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Skin;
RX   MEDLINE=97390076; PubMed=9249015;
RA   Vlasak R., Vilas U., Strobl B., Kreil G.;
RT   "cDNA cloning and expression of secreted Xenopus laevis dipeptidyl
RT   aminopeptidase IV.";
RL   Eur. J. Biochem. 247:107-113(1997).
DR   EMBL; Y08932; CAA70136.1; -; mRNA.
DR   HSSP; P27487; 1PFQ.
DR   MEROPS; S09.003; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004177; F:aminopeptidase activity; IEA.
DR   GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR002471; Pept_S9_AS.
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR002469; Peptidase_S9B.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF00930; DPPIV_N; 1.
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW   Hydrolase; Transmembrane.
SQ   SEQUENCE   748 AA;  85587 MW;  31CCF61B25104E5B CRC64;

  Query Match             53.6%;  Score 2189.5;  DB 2;  Length 748;
  Best Local Similarity   52.0%;  Pred. No. 2.2e-146;
  Matches  396;  Conservative  152;  Mismatches  196;  Indels   17;  Gaps    9;

Qy          1 MKTWVKIVFGVATSAVLALLV-MCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFP 59
              ||||:| :||:   ||:  :| : : |  ::    :  |    ||:|  :  :  |:|
Db          1 MKTWLKWLFGILMGAVVVTVVAVPVALLATKGKGDDRKT---FTLEDYFSDEYRPKSFGL 57

Qy         60 NWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYS 119
               |:|  |:: :  | |::::|::   :  |:|| |: : |:| | || ||::  |: :|
Db         58 KWVSENEFVFRDKD-NVLIFNVDNETTTEIISNTTIHNSNSSFYTLSEDRKYA-LQYNYE 115

Qy        120 KLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDP 179
              ||||:||||:|:|||:   | |  ||||  |||: ||||| |||||::||||:|: ||
Db        116 KLWRHSYTASYHIYDIEKKEIVAANELPNKIQYITWSPVGHKLAYVWENNIYIKEVPGGI 175

Qy        180 PFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYY 239
                || ||  ||| |||||||||||| :| |||||||:   ||| ||||||:||| ||:|
Db        176 STTITTNGEHNKILNGIPDWVYEEEMFSTNYALWWSPDATSLAYVEFNDTDVPVIEYSFY 235

Qy        240 GD--EQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLT 297
              |:  :||| |: |||||||:|| ||:| ::|    | : | |:  |  : | |:| | :
Db        236 GEDSDQYPHTVVIPYPKAGARNPTVRLFAVNTRSLAVINPVEILPPEELRSIDHYISGIN 295

Qy        298 WVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVF 357
              |||| :: :|||:|:||||:|::|     :   |:|        ::| ||| | |  | | |
Db        296 WVTDNKMAVQWLRRIQNVSLLTMC----EGAAWNCQPPV--YKQSSTGWVGYFQPSAPYF 349

Qy        358 SYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEE 417
              : | : |||| |:|:|||||  : : :: : |||| ||  :|  |  : |:| ||  |
Db        350 TQDGLKYYKIISNKEGYKHIHLFEGS-KDPVAITSGNWEVTSIATVASNFLYYVSN--EG 406

Qy        418 YPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLH 477
              :|||| :|:| : |    | :||||: |:|||| |:| ||  :|||:| | |||:|| |::
Db        407 FPGRRQLYKIRLDSGSYSAQCVTCNTRQERCQQYSAYFSKNSKYYSLNCNGPGLPIYTVY 466

Qy        478 DGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLL 537
              :   | | : :|:|::|:  |  ||:| :| | : :|   |||:: ||| ||:|||||||
Db        467 NSSNDNETRTMEDNEDLKLLLDEIQMPTKENKSIIIDGFELWYQLTLPPHFDKSKKYPLL 526

Qy        538 IQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEV 597
              | ||||| || |   | :|| :|||| | :::| :||||: :||||::: :| |||  ||
Db        527 IDVYGGPGSQKVDQFFRLNWATYLASTEKIIVASLDGRGSGYQGDKIMHQIYHKLGTLEV 586

Qy        598 EDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYA 657
```

```
Db        587 QDQITAAKHFSSLGFVDPKRMAIWGWSYGGYVTSMVLGSGSGLFKCGIAVAPVSSWHYYD 646

Qy        658 SVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKAL 717

Db        647 SIYTERYMGLPTPEDNLDNYLSSTVMARAQKFKDVEYLLIHGTADDNVHFQQAAHISKAL 706

Qy        718 VNAQVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSL 758

Db        707 VDAQVDFETMWYTDKDHGIGGTANRHIYTHMSHFLKQCFNI 747


RESULT 14
Q5R7G7_PONPY
ID   Q5R7G7_PONPY PRELIMINARY;      PRT;   765 AA.
AC   Q5R7G7;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein DKFZp469P1419.
GN   Name=DKFZp469P1419;
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RG   The German cDNA Consortium;
RA   Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA   Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR860150; CAH92293.1; -; mRNA.
DR   SMR; Q5R7G7; 38-765.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004177; F:aminopeptidase activity; IEA.
DR   GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR002471; Pept_S9_AS.
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR002469; Peptidase_S9B.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF00930; DPPIV_N; 1.
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW   Hydrolase; Hypothetical protein; Transmembrane.
SQ   SEQUENCE   765 AA;  88038 MW;  6F81ECD98C3D2397 CRC64;

  Query Match           53.6%;  Score 2189;  DB 2;  Length 765;
  Best Local Similarity   51.7%;  Pred. No. 2.5e-146;
  Matches  396;  Conservative  148;  Mismatches  212;  Indels   10;  Gaps    6;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Db          1 MKTPWKVLLGLLGAAALVTIITVPVVLLNKGNDATADSRKTYTLTDYLKNTYRLKLYSLR 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS--NYGLSPDRQFVYLESDY 118

Db         61 WISDHEYLYKQ-ENNILVFNAEYGNSSAFLENSTFDEFGHSINDYSISPDGQFILLEYNY 119

Qy        119 SKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGD 178

Db        120 VKQWRHSYTASYDIYDLNKRQLITEDRIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPNL 179

Qy        179 PPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSY 238

Db        180 PSHRITWTGKEDIIYNGITDWVYEEEVFSAYSALWWSPNGTFLAYAQFNDTEVPLIEYSF 239

Qy        239 YGDE--QYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQ---EVPVPAMIASSDYYF 293

Db        240 YSDESLQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHYL 299

Qy        294 SWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVS 353

Db        300 CDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCLVARQHIEMSTTGWVGRFRPS 359

Qy        354 TPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSN 413

Db        360 EPHFTSDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYISN 419

Qy        414 EFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPI 473

Db        420 EYKGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKYYQLRCSGPGLPL 478

Qy        474 STLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKK 533

Db        479 YTLHSSVNDKGPRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMILPPHFDKSKK 538

Qy        534 YPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLG 593

Db        539 YPLLLDVYAGPCSQKADTVFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRLG 598

Qy        594 VYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSW 653

Db        599 TFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSRW 658

Qy        654 EYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQI 713

Db        659 EYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQQSAQI 718

Qy        714 AKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758

Db        719 SKALVDAGVDFQAMWYTDEDHGIASSTAHQHIYTHMSHFIKQCFSL 764


RESULT 15
DPP4_MOUSE
ID   DPP4_MOUSE     STANDARD;      PRT;   760 AA.
AC   P28843;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
DE   IV) (T-cell activation antigen CD26) (Thymocyte-activating molecule)
DE   (THAM) [Contains: Dipeptidyl peptidase 4 membrane form (Dipeptidyl
DE   peptidase IV membrane form); Dipeptidyl peptidase 4 soluble form
DE   (Dipeptidyl peptidase IV soluble form)].
GN   Name=Dpp4; Synonyms=Cd26;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Swiss; TISSUE=Thymus;
RX   MEDLINE=92129288; PubMed=1370813;
RA   Marguet D.A., Bernard A.-M., Vivier I., Darmoul D., Naquet P.,
RA   Pierres M.;
RT   "cDNA cloning for mouse thymocyte-activating molecule. A
RT   multifunctional ecto-dipeptidyl peptidase IV (CD26) included in a
RT   subgroup of serine proteases.";
RL   J. Biol. Chem. 267:2200-2208(1992).
RN   [2]
RP   SEQUENCE REVISION.
RA   Marguet D.A.;
RL   Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B10.A; TISSUE=Liver;
RX   MEDLINE=95092780; PubMed=7999781;
RA   Bernard A.-M., Mattei M.-G., Pierres M., Marguet D.;
RT   "Structure of the mouse dipeptidyl peptidase IV (CD26) gene.";
```

```
RL   Biochemistry 33:15204-15214(1994).
RN   [4]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [5]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   PROTEIN SEQUENCE OF 1-20.
RX   MEDLINE=91302787; PubMed=1712807;
RA   Vivier I., Marguet D.A., Naquet P., Bonicel J., Black D., Li C.X.-Y.,
RA   Bernard A.-M., Gorvel J.-P., Pierres M.;
RT   "Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl
RT   peptidase IV).";
RL   J. Immunol. 147:447-454(1991).
CC   -!- FUNCTION: Removes N-terminal dipeptides sequentially from
CC       polypeptides having unsubstituted N-termini provided that the
CC       penultimate residue is proline.
CC   -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC       Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC       Zaa is neither Pro nor hydroxyproline.
CC   -!- SUBUNIT: Homodimer.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
CC       soluble form.
CC   -!- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC       by proteolytic processing.
CC   -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; X58384; CAA41274.1; -; mRNA.
DR   EMBL; U12620; AAA82213.1; -; Genomic_DNA.
DR   EMBL; U12599; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12600; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12601; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12602; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12603; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12604; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12605; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12606; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12607; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12608; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12609; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12610; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12611; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12612; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12613; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12614; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12615; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12616; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12617; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12618; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; U12619; AAA82213.1; JOINED; Genomic_DNA.
DR   EMBL; AK085370; BAC39434.1; -; mRNA.
DR   EMBL; BC022183; AAH22183.1; -; mRNA.
DR   HSSP; P27487; 1PFQ.
DR   SMR; P28843; 37-759.
DR   MEROPS; S09.003; -.
DR   Ensembl; ENSMUSG00000035000; Mus musculus.
DR   MGI; MGI:94919; Dpp4.
DR   GO; GO:0016021; C:integral to membrane; TAS.
DR   GO; GO:0046581; C:intercellular canaliculus; IDA.
DR   InterPro; IPR002471; Pept_S9_AS.
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR002469; Peptidase_S9B.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF00930; DPPIV_N; 1.
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW   Aminopeptidase; Direct protein sequencing; Glycoprotein; Hydrolase;
KW   Protease; Serine protease; Signal-anchor; Transmembrane.
FT   CHAIN         1    760       Dipeptidyl peptidase 4 membrane form.
FT   CHAIN        37    760       Dipeptidyl peptidase 4 soluble form (By
FT                                 similarity).
FT   TOPO_DOM      1      6       Cytoplasmic (Potential).
FT   TRANSMEM      7     28       Signal-anchor for type II membrane
FT                                 protein (Potential).
FT   TOPO_DOM     29    760       Extracellular (Potential).
FT   COMPBIAS    295    466       Cys-rich.
FT   ACT_SITE    624    624       Charge relay system (By similarity).
FT   ACT_SITE    702    702       Charge relay system (By similarity).
FT   ACT_SITE    734    734       Charge relay system (By similarity).
FT   CARBOHYD     83     83       N-linked (GlcNAc. . .) (By similarity).
FT   CARBOHYD     90     90       N-linked (GlcNAc. . .) (By similarity).
FT   CARBOHYD    213    213       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    223    223       N-linked (GlcNAc. . .) (By similarity).
FT   CARBOHYD    315    315       N-linked (GlcNAc. . .) (By similarity).
FT   CARBOHYD    328    328       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    514    514       N-linked (GlcNAc. . .) (By similarity).
FT   CARBOHYD    679    679       N-linked (GlcNAc. . .) (By similarity).
FT   DISULFID    379    388       By similarity.
FT   DISULFID    438    441       By similarity.
```

```
FT   DISULFID    448    466       By similarity.
FT   DISULFID    643    756       By similarity.
SQ   SEQUENCE   760 AA;  87437 MW;  A5F644B46E4A3DF8 CRC64;

  Query Match             53.4%;  Score 2183.5;  DB 1;  Length 760;
  Best Local Similarity   52.0%;  Pred. No. 6.1e-146;
  Matches  398;  Conservative  149;  Mismatches  205;  Indels   13;  Gaps    9;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              |||  |:: |:   | |  ::   ::  |:   :  :: |  :| | |  ||  |::
Db          1 MKTPWKVLLGLLGVAALVTIITVPIVLLSK-DEAAADSRRTYSLADYLKSTFRVKSYSLW 59

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              |:|  |||::  :|||:| | | | |   | | | :|     : :|||| || || :| |
Db         60 WVSDFEYLYKQ-ENNILLLNAEHGNSSIFLENSTFESF--GYHSVSPDRLFVLLEYNYVK 116

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
               ||:||||:| |||::  : :   ::|   |:: ||| | |||||::|:||:|  |  |
Db        117 QWRHSYTASYNIYDVNKRQLITEEKIPNNTQWITWSPEGHKLAYVWKNDIYVKVEPHLPS 176

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              :||  | || |:||| |||||||:     |||||||  |||||:|||| :|:| ||:|
Db        177 HRITSTGEENVIYNGITDWVYEEEVFGAYSALWWSPNNTFLAYAQFNDTGVPLIEYSFYS 236

Qy        241 DE--QYPRTINIPYPKAGAKNPVVRIFIID---TTYPAYVGPQEVPVPAMIASSDYYFSW 295
              ||  |||:|: |||||||| || |: ||::    :  :   | ::| || :|  |:|
Db        237 DESLQYPKTVWIPYPKAGAVNPTVKFFIVNIDSLSSSSSAAPIQIPAPASVARGDHYLCD 296

Qy        296 LTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTP 355
              : | |:||: ||||:|:|| ||::|||: :   ||:||  |:|:| | ||| | |  : |
Db        297 VVWATEERISLQWLRRIQNYSVMAICDYDKINLTWNCPSEQQHVEMSTTGWVGRFRPAEP 356

Qy        356 VFSYDAISYYKIFSDKDGYKHI-HYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNE 414
              |: |  |:||| ||||||||| |: ||  ::   || | || |:|  :| | |:| ||:
Db        357 HFTSDGSSFYKIISDKDGYKHICHFPKDK-KDCTFITKGAWEVISIEALTSDYLYYISNQ 415

Qy        415 FEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPIS 474
              ::| || ||:|:| :  :  : ||::| |  |||||  |||  |||| | |:|||:|:
Db        416 YKEMPGGRNLYKIQLTDH-TNVKCLSCDLNPERCQYYAVSFSKEAKYYQLGCWGPGLPLY 474

Qy        475 TLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKY 534
              |||     :|:::||:|  |:  |:::|:| :::  : ::|   ||:||||| ||:||||
Db        475 TLHRSTDHKELRVLEDNSALDRMLQDVQMPSKKLDFIVLNETRFWYQMILPPHFDKSKKY 534

Qy        535 PLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGV 594
              |||: || |||||   : | :|| :|||| | :::|  ||||: :||||:::|: |:||
Db        535 PLLLDVYAGPCSQKADASFRLNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRLGT 594

Qy        595 YEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWE 654
              |||||| | |:|::|||:| ||:|||||||||||:|: | ||:|:|||||||||| ||
Db        595 LEVEDQIEAARQFVKMGFVDSKRVAIWGWSYGGYVTSMVLGSGSGVFKCGIAVAPVSRWE 654

Qy        655 YYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIA 714
              || ||||||:|||| :|||:||:|||||:|||:|: |:|||||||||||||| ||||:
Db        655 YYDSVYTERYMGLPIPEDNLDHYRNSTVMSRAEHFKQVEYLLIHGTADDNVHFQQSAQIS 714

Qy        715 KALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              ||||:| |||||||||:|::||: |  :  |:|:||:|||:|||||
Db        715 KALVDAGVDFQAMWYTDEDHGIASSTAHQHIYSHMSHFLQQCFSL 759


Search completed: January 28, 2006, 00:34:24
Job time : 163 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2006, 00:18:23 ; Search time 136 Seconds
(without alignments)
2455.352 Million cell updates/sec

Title: US-10-681-659-1
Perfect score: 4086
Sequence: 1 MKTWVKIVFGVATSAVLALL.....TNHLYTHMTHFLKQCFSLSD 760

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4086	100.0	760	7	ADN95552	Adn95552 Human BEC
2	4086	100.0	760	8	ADQ21351	Adq21351 Human sof
3	4086	100.0	760	9	AEB94159	Aeb94159 Human wil
4	4069	99.6	760	2	AAW27438	Aaw27438 Human fib
5	4069	99.6	760	6	ABR47452	Abr47452 Breast ca
6	4069	99.6	760	9	ADW14775	Adw14775 Tumor-ass
7	3960	96.9	734	9	AEB94218	Aeb94218 Human sol
8	3901	95.5	723	9	AEB94227	Aeb94227 Human sol
9	3892	95.3	750	9	AEB94161	Aeb94161 Human sol
10	3822.5	93.6	759	2	AAW31963	Aaw31963 Human fib
11	3705.5	90.7	761	9	AEB94163	Aeb94163 Mouse wil
12	2204	53.9	767	6	ABP56699	Abp56699 Rat liver
13	2203	53.9	767	7	ADD46932	Add46932 Rat Prote
14	2199	53.8	767	9	AEB77580	Aeb77580 Rat dipep
15	2198	53.8	767	3	AAB11748	Aab11748 Rat dipep
16	2191.5	53.6	766	7	ADN39272	Adn39272 Cancer/an
17	2190.5	53.6	766	2	AAR40909	Aar40909 Sequence
18	2190.5	53.6	766	5	ABB08991	Abb08991 Human dip
19	2190.5	53.6	766	5	AAG78417	Aag78417 Human dip
20	2190.5	53.6	766	7	ADD27855	Add27855 Human dip
21	2190.5	53.6	766	7	ADD46934	Add46934 Human Pro
22	2190.5	53.6	766	8	ADJ83981	Adj83981 Human ful
23	2190.5	53.6	766	8	ADJ75313	Adj75313 Marker ge
24	2190.5	53.6	766	8	ADO19398	Ado19398 Human PRO
25	2190.5	53.6	766	8	ADO19806	Ado19806 Human PRO
26	2190.5	53.6	766	8	ADO71612	Ado71612 Amino aci
27	2190.5	53.6	766	8	ADO71644	Ado71644 Amino aci
28	2190.5	53.6	766	8	ABM80355	Abm80355 Tumour-as
29	2190.5	53.6	766	8	ADP54458	Adp54458 Human PRO
30	2190.5	53.6	766	8	ADU06688	Adu06688 Novel bro
31	2190.5	53.6	766	8	ADV25525	Adv25525 Human dip
32	2190.5	53.6	766	9	ADY15161	Ady15161 PRO polyp
33	2190.5	53.6	766	9	ADY16580	Ady16580 PRO polyp
34	2190.5	53.6	766	9	ADZ14038	Adz14038 Human dip
35	2190.5	53.6	766	9	AEB94223	Aeb94223 CD26/dipe
36	2187.5	53.5	766	2	AAR54611	Aar54611 Native CD
37	2187.5	53.5	766	5	ABG61910	Abg61910 Prostate
38	2187.5	53.5	766	5	AAO15555	Aao15555 Human dip
39	2187.5	53.5	766	6	ABP56700	Abp56700 Human liv
40	2187.5	53.5	766	7	ADD14045	Add14045 Human src
41	2187.5	53.5	766	7	ADN39604	Adn39604 Cancer/an
42	2187.5	53.5	766	8	ADO19400	Ado19400 Human PRO
43	2185.5	53.5	766	6	ABP55629	Abp55629 Human dpp
44	2185.5	53.5	766	8	ADQ80365	Adq80365 Dipeptidy
45	2185.5	53.5	766	9	AEB77579	Aeb77579 Human dip

ALIGNMENTS

RESULT 1
ADN95552
ID ADN95552 standard; protein; 760 AA.
XX
AC ADN95552;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human BEC/LEC-related protein sequence SeqID475.
XX
KW growth; differentiation; blood endothelial cell; BEC;
KW lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;
KW lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;
KW vasotropic; antiinflammatory; gene therapy; endothelial cell disorder;
KW inflammatory disease; cancer metastasis; lymphatic system; human.
XX
OS Homo sapiens.
XX
PN WO2003080640-A1.
XX
PD 02-OCT-2003.
XX
PF 07-MAR-2003; 2003WO-US006900.
XX
PR 07-MAR-2002; 2002US-0363019P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
PA (LICN) LICENTIA LTD.
XX
PI Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;
XX
DR WPI; 2003-876899/81.
DR N-PSDB; ADN95553.
XX
PS Example 1; SEQ ID NO 475; 176pp; English.
XX
CC This invention relates to a method of differentially modulating the
CC growth or differentiation of blood endothelial cells (BEC) or lymphatic
CC endothelial cells (LEC) comprises contacting endothelial cells with a
CC composition comprising an agent that differentially modulates blood or
CC lymphatic endothelial cells. Treating hereditary lymphoedema comprises
CC identifying a human subject with lymphoedema and with a mutation in at
CC least one allele of a gene encoding a LEC protein, where the mutation
CC correlates with lymphoedema in human subjects, and with the proviso that
CC the LEC protein is not VEGFR-3; and administering to the subject a
CC composition comprising a lymphatic growth agent selected from VEGF-C or
CC VEGF-D polypeptides and polynucleotides. The invention may be useful for

```
CC   the development of compounds with an antiangiogenic, cytostatic,
CC   vasotropic or antiinflammatory activity or for gene therapy. The method
CC   is useful in modulating the growth or differentiation of blood
CC   endothelial cells or lymphatic endothelial cells, in treating hereditary
CC   lymphoedema, in screening for an endothelial cell disorder or
CC   predisposition to the disorder or in monitoring the efficacy or toxicity
CC   of a drug on endothelial cells. The agent is useful in manufacturing a
CC   medicament for the differential modulation of blood vessel endothelial
CC   cell or lymphatic vessel endothelial cell growth or differentiation. The
CC   lymphatic growth agent may also be used in manufacturing a medicament for
CC   the treatment of hereditary lymphoedema resulting from a mutation in a
CC   LEC gene or of other diseases involving the lymphatic vessels, such as
CC   various inflammatory diseases and cancer metastasis via the lymphatic
CC   system. The present sequence is that of a human LEC/BEC differentially
CC   expressed protein which is related to the method of the invention.Note:
CC   This sequence does not appear in the specification but was obtained by
CC   the indexer using the source data given in table 14 of the specification.
XX
SQ   Sequence 760 AA;

  Query Match             100.0%;  Score 4086;  DB 7;  Length 760;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  760;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              ||||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760


RESULT 2
ADQ21351
ID   ADQ21351 standard; protein; 760 AA.
XX
AC   ADQ21351;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated protein - SEQ ID 4171.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 4171; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   protein of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 760 AA;

  Query Match             100.0%;  Score 4086;  DB 8;  Length 760;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  760;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760


RESULT 3
AEB94159
ID   AEB94159 standard; protein; 760 AA.
XX
AC   AEB94159;
XX
DT   06-OCT-2005  (first entry)
XX
DE   Human wild type fibroblast activation protein (FAP) alpha dimer.
XX
KW   immune inhibition; fibroblast activation protein alpha dimer;
KW   FAP alpha dimer; guillain barre syndrome; antiinflammatory; cns-gen.;
KW   immune disorder; neurological disease; autoimmune disease;
KW   immunosuppressive; graft versus host disease; transplant rejection;
KW   endotoxic shock; osteoarthritis; antiarthritic; osteopathic;
KW   musculoskeletal disease; allergy; antiallergic; asthma; antiasthmatic;
KW   inflammation; respiratory disease; atherosclerosis; antiarteriosclerotic;
KW   cardiovascular disease; metabolic disorder; hashimotos disease;
KW   antithyroid; endocrine disease; inflammatory bowel disease;
KW   antiinflammatory; gastrointestinal-gen.; gastrointestinal disease;
KW   rheumatoid arthritis; antirheumatic; multiple sclerosis; neuroprotective;
KW   autoimmune hepatitis; antiinflammatory; hepatotropic;
KW   systemic lupus erythematosus; dermatological; dermatological disease;
KW   uveitis; ophthalmological; autoimmune hemolytic anemia; antianemic;
KW   hematological disease; rheumatic fever; antipyretic; Crohns disease;
KW   psoriasis; antipsoriatic; graves disease; antithyroid;
KW   respiratory syncytial virus infection; respiratory-gen.; virucide;
KW   enzyme.
XX
OS   Homo sapiens.
XX
PN   WO2005071073-A1.
XX
PD   04-AUG-2005.
XX
PF   10-JAN-2005; 2005WO-US000709.
XX
PR   09-JAN-2004; 2004US-0535577P.
XX
PA   (POIN-) POINT THERAPEUTICS INC.
XX
PI   Mclean PA,  Jones B,  Miller GT,  Jesson MI;
XX
DR   WPI; 2005-564220/57.
DR   N-PSDB; AEB94158.
XX
PT   Down-regulating an immune response comprises administering to a subject
PT   in need a fibroblast activation protein (FAP) alpha dimer enzyme in an
PT   amount effective to down-regulate an immune response.
XX
PS   Disclosure; SEQ ID NO 2; 177pp; English.
XX
CC   The invention relates to a method of down-regulating an immune response,
CC   which comprises administering to a subject a fibroblast activation
CC   protein (FAP) alpha dimer enzyme in an amount effective to down-regulate
CC   an immune response. Also included are the following: a composition
CC   comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable
CC   carrier, where the composition is sterile and lacks an adjuvant; a
CC   composition comprising a FAP alpha dimer enzyme in a pharmaceutically
CC   acceptable carrier, and a non-adjuvant second agent; a composition
CC   comprising a FAP alpha dimer enzyme comprising an amino acid substitution
CC   of A657D; and a composition comprising a FAP alpha dimer enzyme lacking
CC   amino acids 269-448 and comprising amino acids 269-448 from mouse FAP.
CC   The method further comprises administering to the subject a second agent.
CC   The second agent is an anti-inflammatory agent, immunosuppressant, or
CC   anti-infective agent such as antibacterial, antiviral, antifungal, anti-
CC   parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild
CC   type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation
CC   mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The
CC   FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a
CC   DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid
CC   substitution relative to wild type FAP alpha dimer. The amino acid
CC   substitution is present in the beta-propeller domain, the catalytic
CC   domain, or an N-linked glycosylation site and alters disulfide bond
CC   formation. The immune response is an especially an IL-I mediated
CC   condition, abnormal immune response selected from inflammation,
CC   autoimmune disease, sepsis, graft versus host disease, transplant
CC   rejection, toxic shock syndrome, allergy, asthma, atherosclerosis,
CC   osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune
CC   response is subsequent to an infection, such as an RSV infection. The
CC   autoimmune disease is selected from c, autoimmune thyroiditis, systemic
CC   lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever,
CC   Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease,
CC   myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple
CC   sclerosis. The subject does not have cancer or a predisposition to
CC   cancer. The present sequence represents the amino acid sequence of human
CC   wild type fibroblast activation protein (FAP) alpha dimer.
XX
SQ   Sequence 760 AA;

  Query Match             100.0%;  Score 4086;  DB 9;  Length 760;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  760;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
```

```
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760


RESULT 4
AAW27438
ID   AAW27438 standard; protein; 760 AA.
XX
AC   AAW27438;
XX
DT   21-JAN-1998  (first entry)
XX
DE   Human fibroblast activation protein-alpha (FAP-alpha).
XX
KW   Fibroblast activation protein-alpha; FAP-alpha; human;
KW   type II integral membrane proteins; epithelial cancer.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Protein         1. .760
FT                   /label= FAP-alpha
FT                   /note= "Fibroblast activation protein alpha"
FT   Region          625. .670
FT                   /note= "alternative sequence found in AAW31963"
XX
PN   WO9734927-A1.
XX
PD   25-SEP-1997.
XX
PF   12-MAR-1997;   97WO-US004215.
XX
PR   18-MAR-1996;   96US-00619280.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Zimmerman R,  Park JE,  Rettig W,  Old LJ;
XX
DR   WPI; 1997-480164/44.
DR   N-PSDB; AAT90148.
XX
PT   Dimeric fibroblast activation protein alpha - capable of degrading
PT   extracellular matrix protein, useful to treat epithelial cancers.
XX
PS   Claim 2; Page 25-28; 40pp; English.
XX
CC   This sequence represents a dimeric fibroblast activation protein alpha
CC   (FAP-alpha) which is capable of degrading extracellular matrix proteins.
CC   The FAP-alpha molecule exhibits structural features typical of type II
CC   integral membrane proteins, including a large COOH-terminal extracellular
CC   domain, a hydrophobic transmembrane segment and a short cytoplasmic tail.
CC   Most common types of epithelial cancers including breast, lung, skin,
CC   pancreas and colorectal carcinomas contain abundant FAP-alpha reactive
CC   stromal fibroblasts which accompany tumour blood vessels. The induction
CC   of FAP-alpha fibroblasts at times and sites of tissue remodelling during
CC   fetal development, tissue repair and carcinogenesis is consistent with
CC   its putative role in normal fibroblast physiology. FAP alpha inhibitors,
CC   particularly a collagen derivative or (S)-valyl-pyrrolidine-2(R)-boronic
CC   acid, may be used to treat various epithelial cancers, e.g. breast, lung,
CC   skin, pancreas or colon cancer. Note: This sequence differs from the
CC   amino acid sequence of FAP-alpha found in AAW31963 between residues 625
CC   to 668
XX
SQ   Sequence 760 AA;

  Query Match             99.6%;  Score 4069;  DB 2;  Length 760;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches  757;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              |||||||||||||||||||||||| |||||||||||||||||| ||||||||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
```

```
Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              ||||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760


RESULT 5
ABR47452
ID   ABR47452 standard; protein; 760 AA.
XX
AC   ABR47452;
XX
DT   12-JUN-2003  (first entry)
XX
DE   Breast cancer associated protein sequence SEQ ID NO:136.
XX
KW   Human; breast cancer; cytostatic; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO2003004989-A2.
XX
PD   16-JAN-2003.
XX
PF   21-JUN-2002; 2002WO-US019669.
XX
PR   21-JUN-2001; 2001US-0299887P.
PR   27-JUN-2001; 2001US-0301572P.
PR   18-JUL-2001; 2001US-0306501P.
PR   25-SEP-2001; 2001US-0325002P.
PR   05-MAR-2002; 2002US-0362585P.
PR   14-MAY-2002; 2002US-0380391P.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Lillie J,  Gannavarapu M,  Glatt K,  Hoersh S,  Kamatkar S;
PI   Mertens M,  Monahan JE,  Myer V,  Wang Y,  Xu Y,  Zhao X,  Meyers RE;
PI   Bast RC,  Hortobagyi GN,  Pusztai L,  Meric F,  Sahin A,  Mills GB;
XX
DR   WPI; 2003-210381/20.
DR   N-PSDB; ACC50144.
XX
PT   Breast cancer diagnosis or treatment by comparing the level of expression
PT   of a marker in a patient sample with that in the control non-breast
PT   cancer sample.
XX
PS   Claim 1; SEQ ID NO 136; 128pp; English.
XX
CC   The present invention describes a method for assessing whether a patient
CC   is afflicted with breast cancer. The method comprises comparing the level
CC   of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC   ABR47386 to ABR47632) in a patient sample and the normal level of
CC   expression of the marker in a control non-breast cancer sample, where a
CC   significant increase in the level of expression of the marker in the
CC   patient sample and the normal level is an indication that the patient is
CC   afflicted with breast cancer. The breast cancer associated sequences from
CC   the present invention have cytostatic activities and can be used in gene
CC   therapy. The method is useful for diagnosing and treating breast cancer.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 760 AA;

  Query Match             99.6%;  Score 4069;  DB 6;  Length 760;
  Best Local Similarity   99.6%;  Pred. No. 0;

  Matches  757;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              |||||||||||||||||||||||| |||||||||||||||||| ||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              ||||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760


RESULT 6
ADW14775
ID   ADW14775 standard; protein; 760 AA.
XX
AC   ADW14775;
XX
DT   07-APR-2005  (first entry)
XX
DE   Tumor-associated antigenic target TAT430.
XX
KW   Tumor-associated antigenic target; TAT430; cancer; neoplasm; cytostatic;
KW   ovary tumor; uterus tumor.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide          1. .28
FT                    /label= Signal_peptide
```

```
FT   Modified-site   10. .15
FT                   /note= "N-myristoylation site"
FT   Protein         29. .760
FT                   /label= Mature_protein
FT   Region          42. .548
FT                   /note= "Dipeptidyl peptidase IV"
FT   Modified-site   49. .52
FT                   /note= "N-glycosylation site"
FT   Modified-site   84. .89
FT                   /note= "N-myristoylation site"
FT   Modified-site   92. .95
FT                   /note= "N-glycosylation site"
FT   Modified-site   99. .102
FT                   /note= "N-glycosylation site"
FT   Modified-site   159. .164
FT                   /note= "N-myristoylation site"
FT   Modified-site   285. .290
FT                   /note= "N-myristoylation site"
FT   Modified-site   314. .317
FT                   /note= "N-glycosylation site"
FT   Modified-site   366. .374
FT                   /note= "Tyrosine kinase phosphorylation site"
FT   Modified-site   402. .410
FT                   /note= "Tyrosine kinase phosphorylation site"
FT   Modified-site   419. .422
FT                   /note= "Amidation site"
FT   Region          438. .443
FT                   /note= "Cytochrome c family heme-binding site signature"
FT   Modified-site   470. .475
FT                   /note= "N-myristoylation site"
FT   Modified-site   542. .547
FT                   /note= "N-myristoylation site"
FT   Region          552. .629
FT                   /note= "Prolyl oligopeptidase family homology"
FT   Modified-site   566. .571
FT                   /note= "N-myristoylation site"
FT   Active-site     599. .629
FT                   /note= "Prolyl endopeptidase family serine active site"
FT   Modified-site   626. .631
FT                   /note= "N-myristoylation site"
FT   Modified-site   636. .639
FT                   /note= "Glycosaminoglycan attachment site"
FT   Modified-site   639. .644
FT                   /note= "N-myristoylation site"
FT   Modified-site   679. .682
FT                   /note= "N-glycosylation site"
FT   Modified-site   738. .743
FT                   /note= "N-myristoylation site"
XX
PN   WO2005003154-A2.
XX
PD   13-JAN-2005.
XX
PF   02-JUL-2004; 2004WO-US021353.
XX
PR   02-JUL-2003; 2003US-0484959P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi A,  Cairns B,  Dowd P,  Frantz G,  French D,  Gonzalez L;
PI   Polakis P,  Smith V,  Wolf B,  Wu TD,  Zhang Z;
XX
DR   WPI; 2005-091764/10.
DR   N-PSDB; ADW14768.
XX
PT   New Tumor-associated Antigenic Target polypeptides useful for diagnosing,
PT   preventing or treating cell proliferative disorders such as cancer or
PT   tumor.
XX
PS   Claim 1; SEQ ID NO 13; 141pp; English.
XX
CC   The present sequence is that of novel tumor-associated antigenic target
CC   polypeptide TAT430. TAT polypeptides have been identified ADW14770-
CC   ADW14776 whose expression is upregulated in particular tumor tissues.
CC   Tissue expression profiling using GeneExpress indicated that TAT430 is
CC   upregulated in breast, pancreatic, colon, adrenal, bladder, kidney,
CC   liver, lung, lymphoid, ovarian, prostate and stomach tumor stroma as
CC   compared to the corresponding healthy tissues. Upregulation in ovarian
CC   tumor tissue was confirmed by quantitative analysis of TAT mRNA
CC   expression, and upregulation in breast, pancreatic, colon, adrenal and
CC   bladder tumors was confirmed by gene expression profiling in silico
CC   (GEPIS) analysis. The TAT polypeptide may comprise the full-length
CC   polypeptide, or may lack the associated signal peptide, or may comprise
CC   an extracellular domain with or without its associated signal peptide, or
CC   may be any of these fused to a heterologous polypeptide such as an
CC   epitope tag, or the Fc region of an immunoglobulin. TAT polypeptides of
CC   the invention, and their encoding nucleic acids, provide targets for the
CC   diagnosis and therapy of cancer.
XX
SQ   Sequence 760 AA;

  Query Match             99.6%;  Score 4069;  DB 9;  Length 760;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches  757;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              |||||||||||||||||||||||||| ||||||||||||||||||||| ||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720

Qy        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||||||||||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
```

RESULT 7
AEB94218
ID AEB94218 standard; protein; 734 AA.
XX
AC AEB94218;
XX
DT 06-OCT-2005 (first entry)
XX
DE Human soluble (FAP) alpha dimer SEQ ID NO:61.
XX
KW immune inhibition; fibroblast activation protein alpha dimer;
KW FAP alpha dimer; guillain barre syndrome; antiinflammatory; cns-gen.;
KW immune disorder; neurological disease; autoimmune disease;
KW immunosuppressive; graft versus host disease; transplant rejection;
KW endotoxic shock; osteoarthritis; antiarthritic; osteopathic;
KW musculoskeletal disease; allergy; antiallergic; asthma; antiasthmatic;
KW inflammation; respiratory disease; atherosclerosis; antiarteriosclerotic;
KW cardiovascular disease; metabolic disorder; hashimotos disease;
KW antithyroid; endocrine disease; inflammatory bowel disease;
KW antiinflammatory; gastrointestinal-gen.; gastrointestinal disease;
KW rheumatoid arthritis; antirheumatic; multiple sclerosis; neuroprotective;
KW autoimmune hepatitis; antiinflammatory; hepatotropic;
KW systemic lupus erythematosus; dermatological; dermatological disease;
KW uveitis; ophthalmological; autoimmune hemolytic anemia; antianemic;
KW hematological disease; rheumatic fever; antipyretic; Crohns disease;
KW psoriasis; antipsoriatic; graves disease; antithyroid;
KW respiratory syncytial virus infection; respiratory-gen.; virucide.
XX
OS Homo sapiens.
XX
PN WO2005071073-A1.
XX
PD 04-AUG-2005.
XX
PF 10-JAN-2005; 2005WO-US000709.
XX
PR 09-JAN-2004; 2004US-0535577P.
XX
PA (POIN-) POINT THERAPEUTICS INC.
XX
PI Mclean PA, Jones B, Miller GT, Jesson MI;
XX
DR WPI; 2005-564220/57.
XX
PT Down-regulating an immune response comprises administering to a subject
PT in need a fibroblast activation protein (FAP) alpha dimer enzyme in an
PT amount effective to down-regulate an immune response.
XX
PS Claim 134; SEQ ID NO 61; 177pp; English.
XX
CC The invention relates to a method of down-regulating an immune response,
CC which comprises administering to a subject a fibroblast activation
CC protein (FAP) alpha dimer enzyme in an amount effective to down-regulate
CC an immune response. Also included are the following: a composition
CC comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable
CC carrier, where the composition is sterile and lacks an adjuvant; a
CC composition comprising a FAP alpha dimer enzyme in a pharmaceutically
CC acceptable carrier, and a non-adjuvant second agent; a composition
CC comprising a FAP alpha dimer enzyme comprising an amino acid substitution
CC of A657D; and a composition comprising a FAP alpha dimer enzyme lacking
CC amino acids 269-448 and comprising amino acids 269-448 from mouse FAP.
CC The method further comprises administering to the subject a second agent.
CC The second agent is an anti-inflammatory agent, immunosuppressant, or
CC anti-infective agent such as antibacterial, antiviral, antifungal, anti-
CC parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild
CC type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation
CC mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The
CC FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a
CC DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid
CC substitution relative to wild type FAP alpha dimer. The amino acid
CC substitution is present in the beta-propeller domain, the catalytic
CC domain, or an N-linked glycosylation site and alters disulfide bond
CC formation. The immune response is an especially an IL-I mediated
CC condition, abnormal immune response selected from inflammation,
CC autoimmune disease, sepsis, graft versus host disease, transplant
CC rejection, toxic shock syndrome, allergy, asthma, atherosclerosis,
CC osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune
CC response is subsequent to an infection, such as an RSV infection. The
CC autoimmune disease is selected from c, autoimmune thyroiditis, systemic
CC lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever,
CC Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease,
CC myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple
CC sclerosis. The subject does not have cancer or a predisposition to
CC cancer. The present sequence represents the amino acid sequence of human
CC soluble fibroblast activation protein (FAP) alpha dimer (corresponding to
CC AEB94159 minus N-terminal amino acids 1-26).
XX
SQ Sequence 734 AA;

Query Match 96.9%; Score 3960; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          27 RPSRVHNSEENTMRALTLKDILNGTFSYKTFFPNWISGQEYLHQSADNNIVLYNIETGQS 86
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 RPSRVHNSEENTMRALTLKDILNGTFSYKTFFPNWISGQEYLHQSADNNIVLYNIETGQS 60

Qy          87 YTILSNRTMKSVNASNYGLSPDRQFVYLESDYSKLWRYSYTATYYIYDLSNGEFVRGNEL 146
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 YTILSNRTMKSVNASNYGLSPDRQFVYLESDYSKLWRYSYTATYYIYDLSNGEFVRGNEL 120

Qy         147 PRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPPFQITFNGRENKIFNGIPDWVYEEEML 206
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 PRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPPFQITFNGRENKIFNGIPDWVYEEEML 180

Qy         207 ATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYGDEQYPRTINIPYPKAGAKNPVVRIFI 266
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 ATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYGDEQYPRTINIPYPKAGAKNPVVRIFI 240

Qy         267 IDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVTDERVCLQWLKRVQNVSVLSICDFRED 326
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 IDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVTDERVCLQWLKRVQNVSVLSICDFRED 300

Qy         327 WQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYDAISYYKIFSDKDGYKHIHYIKDTVEN 386
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 WQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYDAISYYKIFSDKDGYKHIHYIKDTVEN 360

Qy         387 AIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKE 446
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 AIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKE 420

Qy         447 RCQYYTASFSDYAKYYALVCYGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKE 506
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 RCQYYTASFSDYAKYYALVCYGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKE 480

Qy         507 EIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEG 566
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 EIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEG 540

Qy         567 MVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYG 626
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 MVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYG 600

Qy         627 GYVSSLALASGTGLFKCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARA 686
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 GYVSSLALASGTGLFKCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARA 660

Qy         687 EYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLSTNHLYT 746
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 EYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLSTNHLYT 720

Qy         747 HMTHFLKQCFSLSD 760
               ||||||||||||||
Db         721 HMTHFLKQCFSLSD 734
```

```
RESULT 8
AEB94227
ID   AEB94227 standard; protein; 723 AA.
XX
AC   AEB94227;
XX
DT   06-OCT-2005  (first entry)
XX
DE   Human soluble FAP alpha dimer SEQ ID NO:70.
XX
KW   immune inhibition; fibroblast activation protein alpha dimer;
KW   FAP alpha dimer; guillain barre syndrome; antiinflammatory; cns-gen.;
KW   immune disorder; neurological disease; autoimmune disease;
KW   immunosuppressive; graft versus host disease; transplant rejection;
KW   endotoxic shock; osteoarthritis; antiarthritic; osteopathic;
KW   musculoskeletal disease; allergy; antiallergic; asthma; antiasthmatic;
KW   inflammation; respiratory disease; atherosclerosis; antiarteriosclerotic;
KW   cardiovascular disease; metabolic disorder; hashimotos disease;
KW   antithyroid; endocrine disease; inflammatory bowel disease;
KW   antiinflammatory; gastrointestinal-gen.; gastrointestinal disease;
KW   rheumatoid arthritis; antirheumatic; multiple sclerosis; neuroprotective;
KW   autoimmune hepatitis; antiinflammatory; hepatotropic;
KW   systemic lupus erythematosus; dermatological; dermatological disease;
KW   uveitis; ophthalmological; autoimmune hemolytic anemia; antianemic;
KW   hematological disease; rheumatic fever; antipyretic; Crohns disease;
KW   psoriasis; antipsoriatic; graves disease; antithyroid;
KW   respiratory syncytial virus infection; respiratory-gen.; virucide;
KW   CD26 dipeptidyl peptidase IV; DPPIV.
XX
OS   Homo sapiens.
XX
PN   WO2005071073-A1.
XX
PD   04-AUG-2005.
XX
PF   10-JAN-2005; 2005WO-US000709.
XX
PR   09-JAN-2004; 2004US-0535577P.
XX
PA   (POIN-) POINT THERAPEUTICS INC.
XX
PI   Mclean PA,  Jones B,  Miller GT,  Jesson MI;
XX
DR   WPI; 2005-564220/57.
XX
PT   Down-regulating an immune response comprises administering to a subject
PT   in need a fibroblast activation protein (FAP) alpha dimer enzyme in an
PT   amount effective to down-regulate an immune response.
XX
PS   Claim 134; SEQ ID NO 70; 177pp; English.
XX
CC   The invention relates to a method of down-regulating an immune response,
CC   which comprises administering to a subject a fibroblast activation
CC   protein (FAP) alpha dimer enzyme in an amount effective to down-regulate
CC   an immune response. Also included are the following: a composition
CC   comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable
CC   carrier, where the composition is sterile and lacks an adjuvant; a
CC   composition comprising a FAP alpha dimer enzyme in a pharmaceutically
CC   acceptable carrier, and a non-adjuvant second agent; a composition
CC   comprising a FAP alpha dimer enzyme comprising an amino acid substitution
CC   of A657D; and a composition comprising a FAP alpha dimer enzyme lacking
CC   amino acids 269-448 and comprising amino acids 269-448 from mouse FAP.
CC   The method further comprises administering to the subject a second agent.
CC   The second agent is an anti-inflammatory agent, immunosuppressant, or
CC   anti-infective agent such as antibacterial, antiviral, antifungal, anti-
CC   parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild
CC   type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation
CC   mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The
CC   FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a
CC   DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid
CC   substitution relative to wild type FAP alpha dimer. The amino acid
CC   substitution is present in the beta-propeller domain, the catalytic
CC   domain, or an N-linked glycosylation site and alters disulfide bond
CC   formation. The immune response is an especially an IL-I mediated
CC   condition, abnormal immune response selected from inflammation,
CC   autoimmune disease, sepsis, graft versus host disease, transplant
CC   rejection, toxic shock syndrome, allergy, asthma, atherosclerosis,
CC   osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune
CC   response is subsequent to an infection, such as an RSV infection. The
CC   autoimmune disease is selected from c, autoimmune thyroiditis, systemic
CC   lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever,
CC   Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease,
CC   myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple
CC   sclerosis. The subject does not have cancer or a predisposition to
CC   cancer. The present sequence represents the amino acid sequence of human
CC   soluble fibroblast activation protein (FAP) alpha dimer (corresponding to
CC   AEB94159 minus N-terminal amino acids 1-37).
XX
SQ   Sequence 723 AA;

  Query Match             95.5%;  Score 3901;  DB 9;  Length 723;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  723;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         38 TMRALTLKDILNGTFSYKTFFPNWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKS 97
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TMRALTLKDILNGTFSYKTFFPNWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKS 60

Qy         98 VNASNYGLSPDRQFVYLESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSP 157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VNASNYGLSPDRQFVYLESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSP 120

Qy        158 VGSKLAYVYQNNIYLKQRPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPN 217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VGSKLAYVYQNNIYLKQRPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPN 180

Qy        218 GKFLAYAEFNDTDIPVIAYSYYGDEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGP 277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GKFLAYAEFNDTDIPVIAYSYYGDEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGP 240

Qy        278 QEVPVPAMIASSDYYFSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQE 337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 QEVPVPAMIASSDYYFSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQE 300

Qy        338 HIEESRTGWAGGFFVSTPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEA 397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 HIEESRTGWAGGFFVSTPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEA 360

Qy        398 INIFRVTQDSLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSD 457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 INIFRVTQDSLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSD 420

Qy        458 YAKYYALVCYGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEIT 517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 YAKYYALVCYGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEIT 480

Qy        518 LWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGT 577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGT 540

Qy        578 AFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASG 637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASG 600

Qy        638 TGLFKCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLI 697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGLFKCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLI 660

Qy        698 HGTADDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFS 757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 HGTADDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFS 720

Qy        758 LSD 760
              |||
```

Db 721 LSD 723

RESULT 9
AEB94161
ID AEB94161 standard; protein; 750 AA.
XX
AC AEB94161;
XX
DT 06-OCT-2005 (first entry)
XX
DE Human soluble fibroblast activation protein (FAP) alpha dimer.
XX
KW immune inhibition; fibroblast activation protein alpha dimer;
KW FAP alpha dimer; guillain barre syndrome; antiinflammatory; cns-gen.;
KW immune disorder; neurological disease; autoimmune disease;
KW immunosuppressive; graft versus host disease; transplant rejection;
KW endotoxic shock; osteoarthritis; antiarthritic; osteopathic;
KW musculoskeletal disease; allergy; antiallergic; asthma; antiasthmatic;
KW inflammation; respiratory disease; atherosclerosis; antiarteriosclerotic;
KW cardiovascular disease; metabolic disorder; hashimotos disease;
KW antithyroid; endocrine disease; inflammatory bowel disease;
KW antiinflammatory; gastrointestinal-gen.; gastrointestinal disease;
KW rheumatoid arthritis; antirheumatic; multiple sclerosis; neuroprotective;
KW autoimmune hepatitis; antiinflammatory; hepatotropic;
KW systemic lupus erythematosus; dermatological; dermatological disease;
KW uveitis; ophthalmological; autoimmune hemolytic anemia; antianemic;
KW hematological disease; rheumatic fever; antipyretic; Crohns disease;
KW psoriasis; antipsoriatic; graves disease; antithyroid;
KW respiratory syncytial virus infection; respiratory-gen.; virucide;
KW enzyme.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 750
FT /note= "Encoded by taa"
XX
PN WO2005071073-A1.
XX
PD 04-AUG-2005.
XX
PF 10-JAN-2005; 2005WO-US000709.
XX
PR 09-JAN-2004; 2004US-0535577P.
XX
PA (POIN-) POINT THERAPEUTICS INC.
XX
PI Mclean PA, Jones B, Miller GT, Jesson MI;
XX
DR WPI; 2005-564220/57.
DR N-PSDB; AEB94160.
XX
PT Down-regulating an immune response comprises administering to a subject
PT in need a fibroblast activation protein (FAP) alpha dimer enzyme in an
PT amount effective to down-regulate an immune response.
XX
PS Claim 73; SEQ ID NO 4; 177pp; English.
XX
CC The invention relates to a method of down-regulating an immune response,
CC which comprises administering to a subject a fibroblast activation
CC protein (FAP) alpha dimer enzyme in an amount effective to down-regulate
CC an immune response. Also included are the following: a composition
CC comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable
CC carrier, where the composition is sterile and lacks an adjuvant; a
CC composition comprising a FAP alpha dimer enzyme in a pharmaceutically
CC acceptable carrier, and a non-adjuvant second agent; a composition
CC comprising a FAP alpha dimer enzyme comprising an amino acid substitution
CC of A657D; and a composition comprising a FAP alpha dimer enzyme lacking
CC amino acids 269-448 and comprising amino acids 269-448 from mouse FAP.
CC The method further comprises administering to the subject a second agent.
CC The second agent is an anti-inflammatory agent, immunosuppressant, or
CC anti-infective agent such as antibacterial, antiviral, antifungal, anti-
CC parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild
CC type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation
CC mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The
CC FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a
CC DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid
CC substitution relative to wild type FAP alpha dimer. The amino acid
CC substitution is present in the beta-propeller domain, the catalytic
CC domain, or an N-linked glycosylation site and alters disulfide bond
CC formation. The immune response is an especially an IL-I mediated
CC condition, abnormal immune response selected from inflammation,
CC autoimmune disease, sepsis, graft versus host disease, transplant
CC rejection, toxic shock syndrome, allergy, asthma, atherosclerosis,
CC osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune
CC response is subsequent to an infection, such as an RSV infection. The
CC autoimmune disease is selected from c, autoimmune thyroiditis, systemic
CC lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever,
CC Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease,
CC myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple
CC sclerosis. The subject does not have cancer or a predisposition to
CC cancer. The present sequence represents the amino acid sequence of human
CC soluble fibroblast activation protein (FAP) alpha dimer. Note: SEQ ID
CC NO:3 AEB94160 only encodes amino acids 84 to 750 the present sequence.
XX
SQ Sequence 750 AA;

Query Match 95.3%; Score 3892; DB 9; Length 750;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 725; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

```
Qy         17 LALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPNWISGQEYLHQSADNNI 76
Db          6 LLLWVLLLWVPGSTGDAAQPATKRALTLKDILNGTFSYKTFFPNWISGQEYLHQSADNNI 65

Qy         77 VLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSKLWRYSYTATYYIYDLS 136
Db         66 VLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSKLWRYSYTATYYIYDLS 125

Qy        137 NGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPPFQITFNGRENKIFNGI 196
Db        126 NGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPPFQITFNGRENKIFNGI 185

Qy        197 PDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYGDEQYPRTINIPYPKAG 256
Db        186 PDWVYEEEMLATKYALWWSPNGKFLAYAEFNDMDIPVIAYSYYGDEQYPRTINIPYPKAG 245

Qy        257 AKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVTDERVCLQWLKRVQNVS 316
Db        246 AKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVTDERVCLQWLKRVQNVS 305

Qy        317 VLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYDAISYYKIFSDKDGYKH 376
Db        306 VLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYDAISYYKIFSDKDGYKH 365

Qy        377 IHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPGRRNIYRISIGSYPPSK 436
Db        366 IHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPGRRNIYRISIGSYPPSK 425

Qy        437 KCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGRTDQEIKILEENKELEN 496
Db        426 KCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGRTDQEIKILEENKELEN 485

Qy        497 ALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVN 556
Db        486 ALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVN 545

Qy        557 WISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEK 616
Db        546 WISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEK 605

Qy        617 RIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEH 676
Db        606 RIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEH 665
```

```
Qy        677 YKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGL 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 YKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGL 725

Qy        737 SGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||||||||||||
Db        726 SGLSTNHLYTHMTHFLKQCFSLSD 749


RESULT 10
AAW31963
ID   AAW31963 standard; protein; 759 AA.
XX
AC   AAW31963;
XX
DT   21-JAN-1998  (first entry)
XX
DE   Human fibroblast activation protein-alpha (FAP-alpha).
XX
KW   Fibroblast activation protein-alpha; FAP-alpha; human;
KW   type II integral membrane proteins; epithelial cancer.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Protein         1. .759
FT                   /label= FAP-alpha
FT                   /note= "Fibroblast activation protein alpha"
FT   Modified-site   49. .51
FT                   /note= "Asn is N-glycosylated"
FT   Region          191. .209
FT                   /label= fap-1
FT                   /note= "Putative functional element; not defined further
FT                   in specification"
FT   Modified-site   214. .216
FT                   /note= "Asn is N-glycosylated"
FT   Region          220. .240
FT                   /label= fap-2
FT                   /note= "Putative functional element; not defined further
FT                   in specification"
FT   Region          510. .521
FT                   /label= fap-3
FT                   /note= "Putative functional element; not defined further
FT                   in specification"
FT   Region          625. .668
FT                   /note= "alternative sequence found in AAW27438"
FT   Modified-site   637. .639
FT                   /note= "Asn is N-glycosylated"
FT   Modified-site   678. .680
FT                   /note= "Asn is N-glycosylated"
XX
PN   WO9734927-A1.
XX
PD   25-SEP-1997.
XX
PF   12-MAR-1997;   97WO-US004215.
XX
PR   18-MAR-1996;   96US-00619280.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Zimmerman R,  Park JE,  Rettig W,  Old LJ;
XX
DR   WPI; 1997-480164/44.
DR   N-PSDB; AAT89286.
XX
PT   Dimeric fibroblast activation protein alpha - capable of degrading
PT   extracellular matrix protein, useful to treat epithelial cancers.
XX
PS   Example 5; Fig 1; 40pp; English.
XX
CC   This sequence represents the dimeric fibroblast activation protein alpha
CC   (FAP-alpha). The FAP-alpha molecule is capable of degrading extracellular
CC   matrix proteins and exhibits structural features typical of type II
CC   integral membrane proteins, including a large COOH-terminal extracellular
CC   domain, a hydrophobic transmembrane segment and a short cytoplasmic tail.
CC   The putative extracellular domain is described as containing five
CC   potential N-glycosylation sites, eleven conserved cysteine residues and
CC   three segments corresponding to highly conserved catalytic domains
CC   characteristic of serine proteases. Most common types of epithelial
CC   cancers including breast, lung, skin, pancreas and colorectal carcinomas
CC   contain abundant FAP-alpha reactive stromal fibroblasts which accompany
CC   tumour blood vessels. The induction of FAP-alpha fibroblasts at times and
CC   sites of tissue remodelling during foetal development, tissue repair and
CC   carcinogenesis is consistent with its putative role in normal fibroblast
CC   physiology. FAP alpha inhibitors, particularly a collagen derivative or
CC   (S)-valyl-pyrrolidine-2(R)-boronic acid, may be used to treat various
CC   epithelial cancers, e.g. breast, lung, skin, pancreas or colon cancer.
CC   Note: This sequence differs from the amino acid sequence of FAP-alpha
CC   found in AAW27438 between residues 625 to 670
XX
SQ   Sequence 759 AA;

  Query Match             93.6%;  Score 3822.5;  DB 2;  Length 759;
  Best Local Similarity   92.9%;  Pred. No. 0;
  Matches  719;  Conservative    7;  Mismatches   19;  Indels   29;  Gaps    4;

Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              |||||||||||||||||||||||||| |||||||||||||||||| ||||||||||||
Db        181 FQITFNGRENKIFNGIPDWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSRPVFSYD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEY---YA 657
              |||||||||||||||||||||||      :   :|  ::          :| :   |:
Db        601 ITAVRKFIEMGFIDEKRIAIWGWSY-----EIRFITGPCIW----------NWSFQMWYS 645

Qy        658 S-----------VYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVH 706
              |           | | ||||||||||||||||||||||||||||||||||||||||||
Db        646 SGSSLQLGILRVCLHREIHGSPNKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVH 705
```

```
Qy        707 FQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 FQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD 759
```

RESULT 11
AEB94163
ID AEB94163 standard; protein; 761 AA.
XX
AC AEB94163;
XX
DT 06-OCT-2005 (first entry)
XX
DE Mouse wild type fibroblast activation protein (FAP) alpha dimer.
XX
KW immune inhibition; fibroblast activation protein alpha dimer;
KW FAP alpha dimer; guillain barre syndrome; antiinflammatory; cns-gen.;
KW immune disorder; neurological disease; autoimmune disease;
KW immunosuppressive; graft versus host disease; transplant rejection;
KW endotoxic shock; osteoarthritis; antiarthritic; osteopathic;
KW musculoskeletal disease; allergy; antiallergic; asthma; antiasthmatic;
KW inflammation; respiratory disease; atherosclerosis; antiarteriosclerotic;
KW cardiovascular disease; metabolic disorder; hashimotos disease;
KW antithyroid; endocrine disease; inflammatory bowel disease;
KW antiinflammatory; gastrointestinal-gen.; gastrointestinal disease;
KW rheumatoid arthritis; antirheumatic; multiple sclerosis; neuroprotective;
KW autoimmune hepatitis; antiinflammatory; hepatotropic;
KW systemic lupus erythematosus; dermatological; dermatological disease;
KW uveitis; ophthalmological; autoimmune hemolytic anemia; antianemic;
KW hematological disease; rheumatic fever; antipyretic; Crohns disease;
KW psoriasis; antipsoriatic; graves disease; antithyroid;
KW respiratory syncytial virus infection; respiratory-gen.; virucide;
KW enzyme.
XX
OS Mus musculus.
XX
PN WO2005071073-A1.
XX
PD 04-AUG-2005.
XX
PF 10-JAN-2005; 2005WO-US000709.
XX
PR 09-JAN-2004; 2004US-0535577P.
XX
PA (POIN-) POINT THERAPEUTICS INC.
XX
PI Mclean PA, Jones B, Miller GT, Jesson MI;
XX
DR WPI; 2005-564220/57.
DR N-PSDB; AEB94162.
XX
PT Down-regulating an immune response comprises administering to a subject
PT in need a fibroblast activation protein (FAP) alpha dimer enzyme in an
PT amount effective to down-regulate an immune response.
XX
PS Disclosure; SEQ ID NO 6; 177pp; English.
XX
CC The invention relates to a method of down-regulating an immune response,
CC which comprises administering to a subject a fibroblast activation
CC protein (FAP) alpha dimer enzyme in an amount effective to down-regulate
CC an immune response. Also included are the following: a composition
CC comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable
CC carrier, where the composition is sterile and lacks an adjuvant; a
CC composition comprising a FAP alpha dimer enzyme in a pharmaceutically
CC acceptable carrier, and a non-adjuvant second agent; a composition
CC comprising a FAP alpha dimer enzyme comprising an amino acid substitution
CC of A657D; and a composition comprising a FAP alpha dimer enzyme lacking
CC amino acids 269-448 and comprising amino acids 269-448 from mouse FAP.
CC The method further comprises administering to the subject a second agent.
CC The second agent is an anti-inflammatory agent, immunosuppressant, or
CC anti-infective agent such as antibacterial, antiviral, antifungal, anti-
CC parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild
CC type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation
CC mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The
CC FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a
CC DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid
CC substitution relative to wild type FAP alpha dimer. The amino acid
CC substitution is present in the beta-propeller domain, the catalytic
CC domain, or an N-linked glycosylation site and alters disulfide bond
CC formation. The immune response is an especially an IL-I mediated
CC condition, abnormal immune response selected from inflammation,
CC autoimmune disease, sepsis, graft versus host disease, transplant
CC rejection, toxic shock syndrome, allergy, asthma, atherosclerosis,
CC osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune
CC response is subsequent to an infection, such as an RSV infection. The
CC autoimmune disease is selected from c, autoimmune thyroiditis, systemic
CC lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever,
CC Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease,
CC myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple
CC sclerosis. The subject does not have cancer or a predisposition to
CC cancer. The present sequence represents the amino acid sequence of mouse
CC wild type fibroblast activation protein (FAP) alpha dimer.
XX
SQ Sequence 761 AA;

Query Match 90.7%; Score 3705.5; DB 9; Length 761;
Best Local Similarity 89.5%; Pred. No. 7.1e-317;
Matches 681; Conservative 37; Mismatches 42; Indels 1; Gaps 1;

```
Qy          1 MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPN 60
              ||||:| |||| | | |||:|:|||||||||:  | || ||||||||||||||||:|||
Db          1 MKTWLKTVFGVTTLAALALVVICIVLRPSRVYKPEGNTKRALTLKDILNGTFSYKTYFPN 60

Qy         61 WISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNASNYGLSPDRQFVYLESDYSK 120
              ||| ||||||| |:||| ||||| :|| |||| |||||||::||||||||||||||||||
Db         61 WISEQEYLHQSEDDNIVFYNIETRESYIILSNSTMKSVNATDYGLSPDRQFVYLESDYSK 120

Qy        121 LWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180
              |||||||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||
Db        121 LWRYSYTATYYIYDLQNGEFVRGYELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPP 180

Qy        181 FQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYG 240
              ||||: ||||:|||||||||||||||||||||||:|||||| ||||:|||:|||||||
Db        181 FQITYTGRENRIFNGIPDWVYEEEMLATKYALWWSPDGKFLAYVEFNDSDIPIIAYSYYG 240

Qy        241 DEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVT 300
              | |||||||||||||||||||:||:||||| :||| ||||| ||||||||||||||:
Db        241 DGQYPRTINIPYPKAGAKNPVVRVFIVDTTYPHHVGPMEVPVPEMIASSDYYFSWLTWVS 300

Qy        301 DERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYD 360
              |||||||||||||||||||||||||  |:||| |||:||||||||||||||| || |
Db        301 SERVCLQWLKRVQNVSVLSICDFREDWHAWECPKNQEHVEESRTGWAGGFFVSTPAFSQD 360

Qy        361 AISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVTQDSLFYSSNEFEEYPG 420
              | |||||||||||||||||||||||||||||||||||||| |||||||||||||||| |||
Db        361 ATSYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAIYIFRVTQDSLFYSSNEFEGYPG 420

Qy        421 RRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGR 480
              |||||||||: |||||||||||||||||||||||||  ||||||||||:|||||||||
Db        421 RRNIYRISIGNSPPSKKCVTCHLRKERCQYYTASFSYKAKYYALVCYGPGLPISTLHDGR 480

Qy        481 TDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQV 540
              |||||::|||||||||:|:|||||| |||||:   :| ||||||||||||||||||||
Db        481 TDQEIQVLEENKELENSLRNIQLPKVEIKKLKDGGLTFWYKMILPPQFDRSKKYPLLIQV 540

Qy        541 YGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQ 600
              |||||||||:||||||||:|||||||:|||||||||||||||| |:||||||||||||
Db        541 YGGPCSQSVKSVFAVNWITYLASKEGIVIALVDGRGTAFQGDKFLHAVYRKLGVYEVEDQ 600

Qy        601 ITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVY 660
              :|||||||||||||:|||||||||||||||||||||||||||||||||||||||||||:|
Db        601 LTAVRKFIEMGFIDEERIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASIY 660

Qy        661 TERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 SERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHFQNSAQIAKALVNA 720
```

```
Qy        721 QVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSLSD 760
              |||||||||||||||: || | ||||||||||||||||||
Db        721 QVDFQAMWYSDQNHGISSGRSQNHLYTHMTHFLKQCFSLSD 761


RESULT 12
ABP56699
ID   ABP56699 standard; protein; 767 AA.
XX
AC   ABP56699;
XX
DT   27-MAR-2003  (first entry)
XX
DE   Rat liver dipeptidyl peptidase IV (DPP4) protein SEQ ID NO:4.
XX
KW   Tissue-specific endothelial membrane protein; antibacterial; virucide;
KW   fungicide; antiparasitic; anticonvulsant; neuroleptic; cytostatic;
KW   antiparkinsonian; nootropic; neuroprotective; antiasthmatic; infection;
KW   antidiabetic; hypotensive; nephrotropic; antiarthritic; antiinflammatory;
KW   gene therapy; epilepsy; schizophrenia; cancer; Parkinson's disease;
KW   Alzheimer's disease; asthma; diabetes; hypertension; arthritis;
KW   polycystic kidney disease; tissue-specific luminally expressed protein;
KW   inflammatory bowel disease; enzyme.
XX
OS   Rattus norvegicus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 470
FT                   /note= "encoded by CAG"
XX
PN   WO2002100336-A2.
XX
PD   19-DEC-2002.
XX
PF   07-JUN-2002; 2002WO-US018185.
XX
PR   08-JUN-2001; 2001US-0297021P.
PR   12-JUL-2001; 2001US-0305117P.
XX
PA   (TARG-) TARGET PROTEIN TECHNOLOGIES INC.
XX
PI   Roben PW,  Stevens AC;
XX
DR   WPI; 2003-167367/16.
DR   N-PSDB; ABZ22556.
XX
PT   Delivering a therapeutic agent (e.g. immunosuppressant) to a specific
PT   tissue, comprises administering an amount of a therapeutic complex
PT   comprising a therapeutic group, a ligand that binds to a tissue-specific
PT   protein, and a linker.
XX
PS   Claim 24; Page 99-100; 123pp; English.
XX
CC   The present invention describes a method (M1) for delivering a
CC   therapeutic agent to a specific tissue, comprising administering an
CC   amount of a therapeutic complex comprising a ligand which binds to a
CC   tissue-specific luminally expressed protein, a therapeutic group, and a
CC   linker which links the group to the ligand. Tissue-specific endothelial
CC   membrane proteins from the present invention can have antibacterial,
CC   virucide, fungicide, antiparasitic, anticonvulsant, neuroleptic,
CC   cytostatic, antiparkinsonian, nootropic, neuroprotective, antiasthmatic,
CC   antidiabetic, hypotensive, nephrotropic, antiarthritic and
CC   antiinflammatory activities, and can be used in gene therapy. (M1) is
CC   useful in targeting pharmaceuticals or other therapeutics to specific
CC   tissues using tissue-specific endothelial membrane proteins. A
CC   therapeutic complex may be used to treat or diagnose any disease for
CC   which a tissue- or organ-specific treatment would be efficacious, such as
CC   in cases of infections (e.g. bacterial, viral, fungal and parasitic),
CC   epilepsy, schizophrenia, cancer, Parkinson's disease, Alzheimer's
CC   disease, asthma, diabetes, hypertension, polycystic kidney disease,
CC   arthritis, and inflammatory bowel disease. The present sequence
CC   represents a rat liver dipeptidyl peptidase IV (DPP4), which is used in
CC   an example from the present invention
XX
SQ   Sequence 767 AA;

  Query Match             53.9%;  Score 2204;  DB 6;  Length 767;
  Best Local Similarity   52.4%;  Pred. No. 1.8e-184;
  Matches  405;  Conservative  148;  Mismatches  198;  Indels   22;  Gaps   11;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEE---NTMRALTLKDILNGTFSYKT 56
              |||  |:: | :  :|:: :: : :||      | :|   :: |  || | |  ||  |:
Db          1 MKTPWKVLLGLLGVAALVTIITVPVVLL-----NKDEAAADSARTYTLADYLKNTFRVKS 55

Qy         57 FFPNWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMK--SVNASNYGLSPDRQFVYL 114
              :  |:|  |||::  :|||:|:| | | |   | | | :    : |:| :|||| || |
Db         56 YSLRWVSDSEYLYKQ-ENNILLFNAEHGNSSIFLENSTFEIFGDSISDYSVSPDRLFVLL 114

Qy        115 ESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQ 174
              | :| | ||:||||:| ||||:  : :   ::|   |:: ||  | |||||:|:||:|
Db        115 EYNYVKQWRHSYTASYSIYDLNKRQLITEEKIPNNTQWITWSQEGHKLAYVWKNDIYVKI 174

Qy        175 RPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVI 234
              |  |  :||  |:|| ||||| ||||||||:     |||||||| |||||:|||| :|:|
Db        175 EPHLPSHRITSTGKENVIFNGINDWVYEEEIFGAYSALWWSPNGTFLAYAQFNDTGVPLI 234

Qy        235 AYSYYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDT---TYPAYVGPQEVPVPAMIASS 289
              ||:| ||  |||:|: |||||||| || |: ||::|   :       | ::  || : :
Db        235 EYSFYSDESLQYPKTVWIPYPKAGAVNPTVKFFIVNTDSLSSTTTTIPMQITAPASVTTG 294

Qy        290 DYYFSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGG 349
              |:|   : ||:::|: ||||:|:|| ||::|||: :    |:|| |||||| | ||| |
Db        295 DHYLCDVAWVSEDRISLQWLRRIQNYSVMAICDYDKTTLVWNCPTTQEHIETSATGWCGR 354

Qy        350 FFVSTPVFSYDAISYYKIFSDKDGYKHI-HYIKDTVENAI--QITSGKWEAINIFRVTQD 406
              |  : | |: |  |:||| |||||||||  : ||     :   || | || |:|  :| |
Db        355 FRPAEPHFTSDGSSFYKIVSDKDGYKHICQFQKDRKPEQVCTFITKGAWEVISIEALTSD 414

Qy        407 SLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVC 466
              |:| |||::| || ||:|:| :  :  :|||::| |  ||||||: | |  |||| | |
Db        415 YLYYISNEYKEMPGGRNLYKIQLTDH-TNKKCLSCDLNPERCQYYSVSLSKEAKYYGLGC 473

Qy        467 YGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPP 526
              |||:|: |||     :|:::||:|  |:  |:::|:| :::  : ::|   ||:|||||
Db        474 RGPGLPLYTLHRSTDQKELRVLEDNSALDKMLQDVQMPSKKLDFIVLNETRFWYQMILPP 533

Qy        527 QFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLY 586
              ||:|||||||| || |||||   : | :|| :|||| | :::|  ||||: :||||:::
Db        534 HFDKSKKYPLLIDVYAGPCSQKADAAFRLNWATYLASTENIIVASFDGRGSGYQGDKIMH 593

Qy        587 AVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIA 646
              |: ::||  |||||| | |:|::|||:| ||:|||||||||||:|: | ||:|:||||||
Db        594 AINKRLGTLEVEDQIEAARQFLKMGFVDSKRVAIWGWSYGGYVTSMVLGSGSGVFKCGIA 653

Qy        647 VAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVH 706
              ||||| |||| ||||||:||||| :|||:||:|||||:||| |: |:||||||||||||
Db        654 VAPVSRWEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVH 713

Qy        707 FQNSAQIAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              || ||||:||||:| ||||||||:|::||: |  :  |:|:||:|||:|||||
Db        714 FQQSAQISKALVDAGVDFQAMWYTDEDHGIASSTAHQHIYSHMSHFLQQCFSL 766


RESULT 13
ADD46932
ID   ADD46932 standard; protein; 767 AA.
XX
AC   ADD46932;
XX
DT   02-DEC-2004  (revised)
DT   29-JAN-2004  (first entry)
XX
DE   Rat Protein AAA41096, SEQ ID NO 12618.
XX
KW   Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
```

```
KW   chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS   Rattus norvegicus.
OS   Unidentified.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; AAA41096.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Example 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a rat protein (described in Table 3
CC   of the specification) which is differentially expressed during pain.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 767 AA;

  Query Match             53.9%;  Score 2203;  DB 7;  Length 767;
  Best Local Similarity   52.4%;  Pred. No. 2.2e-184;
  Matches  405;  Conservative  148;  Mismatches  198;  Indels   22;  Gaps   11;

Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEE---NTMRALTLKDILNGTFSYKT 56
              |||  |:: | :  :|:: :: : :||      | :|   :: |  || | |  ||  |:
Db          1 MKTPWKVLLGLLGVAALVTIITVPVVLL-----NKDEAAADSARTYTLADYLKNTFRVKS 55

Qy         57 FFPNWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMK--SVNASNYGLSPDRQFVYL 114
              :   |:|  |||::  :|||:|:| | | |   | | | :   : |:| :|||| || |
Db         56 YSLRWVSDSEYLYKQ-ENNILLFNAEHGNSSIFLENSTFEIFGDSISDYSVSPDRLFVLL 114

Qy        115 ESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQ 174
              | :| | ||:||||:| ||||:  : :   ::|   |:: ||  | |||||::|:||:|
Db        115 EYNYVKQWRHSYTASYSIYDLNKRQLITEEKIPNNTQWITWSQEGHKLAYVWKNDIYVKI 174

Qy        175 RPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVI 234
              |  |  :||  |:|| ||||| |||||||:     |||||||| |||||:|||| :|:|
Db        175 EPHLPSHRITSTGKENVIFNGINDWVYEEEIFGAYSALWWSPNGTFLAYAQFNDTGVPLI 234

Qy        235 AYSYYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDT---TYPAYVGPQEVPVPAMIASS 289
              ||:| ||  |||:|: |||||||| || |: ||::|   :      | ::  || : :
Db        235 EYSFYSDESLQYPKTVWIPYPKAGAVNPTVKFFIVNTDSLSSTTTTIPMQITAPASVTTG 294

Qy        290 DYYFSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGG 349
              |:|   : ||:::|: ||||:|:|| ||::|||: :    |:|| |||||| | ||| |
Db        295 DHYLCDVAWVSEDRISLQWLRRIQNYSVMAICDYDKTTLVWNCPTTQEHIETSATGWCGR 354

Qy        350 FFVSTPVFSYDAISYYKIFSDKDGYKHI-HYIKDTVENAI--QITSGKWEAINIFRVTQD 406
              |  : | |: |  |:||| |||||||||  : ||     :   || | || |:|  :| |
Db        355 FRPAEPHFTSDGSSFYKIVSDKDGYKHICQFQKDRKPEQVCTFITKGAWEVISIEALTSD 414

Qy        407 SLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVC 466
              |:| |||::| || ||:|:| :  :  :|||::| |  ||||||: | |  |||| | |
Db        415 YLYYISNEYKEMPGGRNLYKIQLTDH-TNKKCLSCDLNPERCQYYSVSLSKEAKYYQLGC 473

Qy        467 YGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPP 526
              |||:|: |||      :|:::||:|  |:  |:::|:| :::  : ::|   ||:|||||
Db        474 RGPGLPLYTLHRSTDQKELRVLEDNSALDKMLQDVQMPSKKLDFIVLNETRFWYQMILPP 533

Qy        527 QFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLY 586
              ||:|||||||| || |||||   : | :|| :|||| | :::|  ||||: :||||:::
Db        534 HFDKSKKYPLLIDVYAGPCSQKADAAFRLNWATYLASTENIIVASFDGRGSGYQGDKIMH 593

Qy        587 AVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIA 646
              |: ::||  |||||| | |:|::|||:| ||:|||||||||||:|: | ||:|:||||||
Db        594 AINKRLGTLEVEDQIEAARQFLKMGFVDSKRVAIWGWSYGGYVTSMVLGSGSGVFKCGIA 653

Qy        647 VAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVH 706
              ||||| |||| |||||||:||||| :|||:||:|||||:||| |: |:|||||||||||||
Db        654 VAPVSRWEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVH 713

Qy        707 FQNSAQIAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              || ||||:||||:| ||||||||:|::||: |  :  |:|:||:|||:|||||
Db        714 FQQSAQISKALVDAGVDFQAMWYTDEDHGIASSTAHQHIYSHMSHFLQQCFSL 766


RESULT 14
AEB77580
ID   AEB77580 standard; protein; 767 AA.
XX
AC   AEB77580;
XX
DT   06-OCT-2005  (first entry)
XX
DE   Rat dipeptidyl peptidase IV enzyme - SEQ ID 2.
XX
KW   autism; nootropic; asperger syndrome; enzyme; dipeptidyl peptidase IV.
XX
OS   Rattus norvegicus.
XX
PN   US2005170333-A1.
XX
PD   04-AUG-2005.
XX
PF   03-FEB-2004; 2004US-00770712.
XX
PR   03-FEB-2004; 2004US-00770712.
XX
PA   (VOJD/) VOJDANI A.
XX
PI   Vojdani A;
XX
DR   WPI; 2005-562713/57.
XX
PT   Determining etiology of autistic spectrum disorder in patient, by
PT   determining level of infectious agent/toxic chemical/dietary protein
PT   derived antigen in samples of patient, comparing it with normal level of
```

PT antigens of control subjects.
XX
PS Disclosure; SEQ ID NO 2; 89pp; English.
XX
CC The invention comprises a method of determining etiology of an autistic
CC spectrum disorder in a patient. The method involves determining the level
CC of an infectious agent, toxic chemical, or dietary protein derived
CC antigen, or their antibodies in samples of patient, and comparing
CC antigens/antibodies levels with normal levels of antigens/antibodies from
CC control subjects. The method of the invention is useful for determining
CC the etiology of an autistic spectrum disorder, such as autism, pervasive
CC development disorder and Asperger's syndrome. The present amino acid
CC sequence represents a rat dipeptidyl peptidase IV enzyme that was used in
CC the exemplification of the invention.
XX
SQ Sequence 767 AA;

Query Match 53.8%; Score 2199; DB 9; Length 767;
Best Local Similarity 52.3%; Pred. No. 4.9e-184;
Matches 404; Conservative 149; Mismatches 198; Indels 22; Gaps 11;

```
Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEE---NTMRALTLKDILNGTFSYKT 56
              |||  |:: | :  :|:: :: : :||      | :|   :: |  || | |  ||  |:
Db          1 MKTPWKVLLGLLGVAALVTIITVPVVLL-----NKDEAAADSARTYTLADYLKNTFRVKS 55

Qy         57 FFPNWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMK--SVNASNYGLSPDRQFVYL 114
              :  |:|  |||::  :|||:|:| | | |   | | | :    : |:| :|||| || |
Db         56 YSLRWVSDSEYLYKQ-ENNILLFNAEHGNSSIFLENSTFEIFGDSISDYSVSPDRLFVLL 114

Qy        115 ESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQ 174
              | :| | ||:||||:| ||||:  : :   ::|   |:: ||  | |||||::|:||:|
Db        115 EYNYVKQWRHSYTASYSIYDLNKRQLITEEKIPNNTQWITWSQEGHKLAYVWKNDIYVKI 174

Qy        175 RPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVI 234
              |  |  :||  |:|| ||||| |||||||:     |||||||| |||||:|||| :|:|
Db        175 EPHLPSHRITSTGKENVIFNGINDWVYEEEIFGAYSALWWSPNGTFLAYAQFNDTGVPLI 234

Qy        235 AYSYYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDT---TYPAYVGPQEVPVPAMIASS 289
              ||:| ||  |||:|: |||||||| || |: ||::|    :      | ::  || : :
Db        235 EYSFYSDESLQYPKTVWIPYPKAGAVNPTVKFFIVNTDSLSSTTTTIPMQITAPASVTTG 294

Qy        290 DYYFSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGG 349
              |:|   : ||:::|: ||||:|:|| ||::|||: :    |:|| |:|||| | ||| |
Db        295 DHYLCDVAWVSEDRISLQWLRRIQNYSVMAICDYDKTTLVWNCPTTREHIETSATGWCGR 354

Qy        350 FFVSTPVFSYDAISYYKIFSDKDGYKHI-HYIKDTVENAI--QITSGKWEAINIFRVTQD 406
              |  : | |: |  |:||| |||||||| : ||     :   || | || |:|  :| |
Db        355 FRPAEPHFTSDGSSFYKIVSDKDGYKHICQFQKDRKPEQVCTFITKGAWEVISIEALTSD 414

Qy        407 SLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVC 466
              |:| |||::| || ||:|:| :  :  :|||::| |  ||||||: | |  |||| | |
Db        415 YLYYISNEYKEMPGGRNLYKIQLTDH-TNKKCLSCDLNPERCQYYSVSLSKEAKYYQLGC 473

Qy        467 YGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPP 526
              |||:|: |||     :|:::||:|  |:  |:::|:| :::  : ::|   ||:|||||
Db        474 RGPGLPLYTLHRSTDQKELRVLEDNSALDKMLQDVQMPSKKLDFIVLNETRFWYQMILPP 533

Qy        527 QFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLY 586
              ||:|||||||| || |||||   : | :|| :|||| | :::|  ||||: :||||:::
Db        534 HFDKSKKYPLLIDVYAGPCSQKADAAFRLNWATYLASTENIIVASFDGRGSGYQGDKIMH 593

Qy        587 AVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIA 646
              |: ::|| |||||| | |:|::|||:| ||:|||||||||||:|: | ||:|:||||||
Db        594 AINKRLGTLEVEDQIEAARQFLKMGFVDSKRVAIWGWSYGGYVTSMVLGSGSGVFKCGIA 653

Qy        647 VAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVH 706
              ||||| |||| ||||||:|||||  :|||:||:|||||:||| |: |:|||||||||||||
Db        654 VAPVSRWEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVH 713

Qy        707 FQNSAQIAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
              || ||||:||||:| ||||||||:|::||: |  :  |:|:||:|||:|||||
Db        714 FQQSAQISKALVDAGVDFQAMWYTDEDHGIASSTAHQHIYSHMSHFLQQCFSL 766
```

RESULT 15
AAB11748
ID AAB11748 standard; protein; 767 AA.
XX
AC AAB11748;
XX
DT 28-OCT-2000 (first entry)
XX
DE Rat dipeptidyl peptidase IV (DPPIV).
XX
KW Dipeptidyl peptidase IV; DPPIV; rat; glomerulus;
KW F-16 monoclonal antibody; complement-mediated cytotoxicity; cell damage;
KW myocardial infarction; reperfusion injury; multiple sclerosis; nephritis;
KW neuropathy; arthritis; cardiant; neuroprotective; vasotropic;
KW antiarthritic.
XX
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT Region 35. .767
FT /note= "This region is specifically claimed"
XX
PN JP2000143699-A.
XX
PD 26-MAY-2000.
XX
PF 11-NOV-1998; 98JP-00321079.
XX
PR 11-NOV-1998; 98JP-00321079.
XX
PA (SHIO) SHIONOGI & CO LTD.
XX
DR WPI; 2000-445041/39.
XX
PT Monoclonal antibody specific for dipeptidyl peptidase, useful for
PT prevention and treatment of nephritis, arthritis, neuropathy and
PT myocardial infarction resulting from complement-dependent cell damage.
XX
PS Claim 2; Page 7-9; 11pp; Japanese.
XX
CC This sequence represents rat dipeptidyl peptidase IV (DPPIV). The
CC invention relates to the murine monoclonal antibody F-16 (FERMP-17016),
CC which suppresses complement-mediated cell damage by specifically binding
CC to DPPIV, thereby inhibiting complement activation. The invention also
CC relates to the F-16 hybridoma which secretes the monoclonal antibody and
CC a method of screening complement suppressive agents using the antibody
CC and DPPIV. The F-16 hybridoma was produced by fusion of a mouse myeloma
CC cell and a splenocyte obtained from a mouse inoculated with rat
CC glomerulus protein extract. The F-16 antibody was subsequently found to
CC bind to DPPIV. The F-16 antibody is useful for treating diseases caused
CC by complement-mediated cytotoxicity, such as myocardial infarction,
CC reperfusion injury, multiple sclerosis, nephritis, neuropathy and
CC arthritis. The antibody is also useful for screening agents with
CC potential complement suppression activity
XX
SQ Sequence 767 AA;

Query Match 53.8%; Score 2198; DB 3; Length 767;
Best Local Similarity 52.3%; Pred. No. 6e-184;
Matches 404; Conservative 149; Mismatches 198; Indels 22; Gaps 11;

```
Qy          1 MKTWVKIVFG-VATSAVLALLVMCIVLRPSRVHNSEE---NTMRALTLKDILNGTFSYKT 56
              |||  |:: | :  :|:: :: : :||      | :|   :: |  || | |  ||  |:
Db          1 MKTPWKVLLGLLGVAALVTIITVPVVLL-----NKDEAAADSRRTYTLADYLKNTFRVKS 55

Qy         57 FFPNWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMK--SVNASNYGLSPDRQFVYL 114
              :  |:|  |||::  :|||:|:| | | |   | | | :    : |:| :|||| || |
Db         56 YSLRWVSDSEYLYKQ-ENNILLFNAEHGNSSIFLENSTFEIFGDSISDYSVSPDRLFVLL 114

Qy        115 ESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQ 174
              | :| | ||:||||:| ||||:  : :   ::|   |:: ||  | |||||::|:||:|
Db        115 EYNYVKQWRHSYTASYSIYDLNKRQLITEEKIPNNTQWITWSQEGHKLAYVWKNDIYVKI 174
```

```
Qy        175 RPGDPPFQITFNGRENKIFNGIPDWVYEEEMLATKYALWWSPNGKFLAYAEFNDTDIPVI 234
Db        175 EPHLPSHRITSTGKENVIFNGINDWVYEEEIFGAYSALWWSPNGTFLAYAQFNDTGVPLI 234

Qy        235 AYSYYGDE--QYPRTINIPYPKAGAKNPVVRIFIIDT---TYPAYVGPQEVPVPAMIASS 289
Db        235 EYSFYSDESLQYPKTVWIPYPKAGAVNPTVKFFIVNTDSLSSTTTTIPMQITAPASVTTG 294

Qy        290 DYYFSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGG 349
Db        295 DHYLCDVAWVSEDRISLQWLRRIQNYSVMAICDYDKTNLVWNCPTTQEHIETSATGWVGR 354

Qy        350 FFVSTPVFSYDAISYYKIFSDKDGYKHI-HYIKDTVENAI--QITSGKWEAINIFRVTQD 406
Db        355 FRPAEPHFTSDGSSFYKIVSDKDGYKHICEFQKDRKPEQVCTFITKGAWEVISIEALTSD 414

Qy        407 SLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVC 466
Db        415 YLYYISNEYKEMPGGRNLYKIQLTDH-TNKKCLSCDLNPERCQYYSVSLSKEAKYYQLGC 473

Qy        467 YGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPKEEIKKLEVDEITLWYKMILPP 526
Db        474 RGPGLPLYTLHRSTDQKELRVLEDNSALDKMLQDVQMPSKKLDFIVLNETRFWYQMILPP 533

Qy        527 QFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLY 586
Db        534 HFDKSKKYPLLIDVYAGPCSQKADAAFRLNWATYLASTENIIVASFDGRGSGYQGDKIMH 593

Qy        587 AVYRKLGVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIA 646
Db        594 AINKRLGTLEVEDEIEAARQFLKMGFVDSKRVAIWGWSYGGYVTSMVLGSGSGVFKCGIA 653

Qy        647 VAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVH 706
Db        654 VAPVSRWEYYDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVH 713

Qy        707 FQNSAQIAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLYTHMTHFLKQCFSL 758
Db        714 FQQSAQISKALVDAGVDFQAMWYTDEDHGIASSTAHQHIYSHMSHFLQQCFSL 766
```

Search completed: January 28, 2006, 00:31:33
Job time : 141 secs

This Page Blank (uspto)